Cochrane Library

Cochrane Database of Systematic Reviews

Music therapy for autistic people (Review)

Geretsegger M, Fusar-Poli L, Elefant C, Mössler KA, Vitale G, Gold C

Geretsegger M, Fusar-Poli L, Elefant C, Mössler KA, Vitale G, Gold C.
Music therapy for autistic people.
Cochrane Database of Systematic Reviews 2022, Issue 5. Art. No.: CD004381.
DOI: 10.1002/14651858.CD004381.pub4.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Music therapy for autistic people

Monika Geretsegger[1], Laura Fusar-Poli[2], Cochavit Elefant[3], Karin A Mössler[1], Giovanni Vitale[4], Christian Gold[1,5]

[1]GAMUT - The Grieg Academy Music Therapy Research Centre, NORCE Norwegian Research Centre, Bergen, Norway. [2]Department of Clinical and Experimental Medicine, Psychiatry Unit, University of Catania, Catania, Italy. [3]Department of Creative Arts Therapies, University of Haifa, Haifa, Israel. [4]Child and Adolescent Neuropsychiatry, University of Naples Federico II, Naples, Italy. [5]Department of Clinical and Health Psychology, University of Vienna, Vienna, Austria

Contact: Christian Gold, chgo@norceresearch.no.

Editorial group: Cochrane Developmental, Psychosocial and Learning Problems Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 5, 2022.

Citation: Geretsegger M, Fusar-Poli L, Elefant C, Mössler KA, Vitale G, Gold C.Music therapy for autistic people. *Cochrane Database of Systematic Reviews* 2022, Issue 5. Art. No.: CD004381. DOI: 10.1002/14651858.CD004381.pub4.

ABSTRACT

Background

Social interaction and social communication are among the central areas of difficulty for autistic people. Music therapy uses music experiences and the relationships that develop through them to enable communication and expression, thus attempting to address some of the core problems of autistic people. Music therapy has been applied in autism since the early 1950s, but its availability to autistic individuals varies across countries and settings. The application of music therapy requires specialised academic and clinical training which enables therapists to tailor the intervention to the specific needs of the individual. The present version of this review on music therapy for autistic people is an update of the previous Cochrane review update published in 2014 (following the original Cochrane review published in 2006).

Objectives

To review the effects of music therapy, or music therapy added to standard care, for autistic people.

Search methods

In August 2021, we searched CENTRAL, MEDLINE, Embase, eleven other databases and two trials registers. We also ran citation searches, checked reference lists, and contacted study authors to identify additional studies.

Selection criteria

All randomised controlled trials (RCTs), quasi-randomised trials and controlled clinical trials comparing music therapy (or music therapy alongside standard care) to 'placebo' therapy, no treatment, or standard care for people with a diagnosis of autism spectrum disorder were considered for inclusion.

Data collection and analysis

We used standard Cochrane methodological procedures. Four authors independently selected studies and extracted data from all included studies. We synthesised the results of included studies in meta-analyses. Four authors independently assessed risk of bias (RoB) of each included study using the original RoB tool as well as the certainty of evidence using GRADE.

Main results

We included 16 new studies in this update which brought the total number of included studies to 26 (1165 participants). These studies examined the short- and medium-term effect of music therapy (intervention duration: three days to eight months) for autistic people in individual or group settings. More than half of the studies were conducted in North America or Asia. Twenty-one studies included children

aged from two to 12 years. Five studies included children and adolescents, and/or young adults. Severity levels, language skills, and cognition were widely variable across studies.

Measured immediately post-intervention, music therapy compared with 'placebo' therapy or standard care was more likely to positively effect global improvement (risk ratio (RR) 1.22, 95% confidence interval (CI) 1.06 to 1.40; 8 studies, 583 participants; moderate-certainty evidence; number needed to treat for an additional beneficial outcome (NNTB) = 11 for low-risk population, 95% CI 6 to 39; NNTB = 6 for high-risk population, 95% CI 3 to 21) and to slightly increase quality of life (SMD 0.28, 95% CI 0.06 to 0.49; 3 RCTs, 340 participants; moderate-certainty evidence, small to medium effect size). In addition, music therapy probably results in a large reduction in total autism symptom severity (SMD −0.83, 95% CI −1.41 to −0.24; 9 studies, 575 participants; moderate-certainty evidence). No clear evidence of a difference between music therapy and comparison groups at immediately post-intervention was found for social interaction (SMD 0.26, 95% CI −0.05 to 0.57, 12 studies, 603 participants; low-certainty evidence); non-verbal communication (SMD 0.26, 95% CI −0.03 to 0.55; 7 RCTs, 192 participants; low-certainty evidence); and verbal communication (SMD 0.30, 95% CI −0.18 to 0.78; 8 studies, 276 participants; very low-certainty evidence). Two studies investigated adverse events with one (36 participants) reporting no adverse events; the other study found no differences between music therapy and standard care immediately post-intervention (RR 1.52, 95% CI 0.39 to 5.94; 1 study, 290 participants; moderate-certainty evidence).

Authors' conclusions

The findings of this updated review provide evidence that music therapy is probably associated with an increased chance of global improvement for autistic people, likely helps them to improve total autism severity and quality of life, and probably does not increase adverse events immediately after the intervention. The certainty of the evidence was rated as 'moderate' for these four outcomes, meaning that we are moderately confident in the effect estimate. No clear evidence of a difference was found for social interaction, non-verbal communication, and verbal communication measured immediately post-intervention. For these outcomes, the certainty of the evidence was rated as 'low' or 'very low', meaning that the true effect may be substantially different from these results. Compared with earlier versions of this review, the new studies included in this update helped to increase the certainty and applicability of this review's findings through larger sample sizes, extended age groups, longer periods of intervention and inclusion of follow-up assessments, and by predominantly using validated scales measuring generalised behaviour (i.e. behaviour outside of the therapy context). This new evidence is important for autistic individuals and their families as well as for policymakers, service providers and clinicians, to help in decisions around the types and amount of intervention that should be provided and in the planning of resources. The applicability of the findings is still limited to the age groups included in the studies, and no direct conclusions can be drawn about music therapy in autistic individuals above the young adult age. More research using rigorous designs, relevant outcome measures, and longer-term follow-up periods is needed to corroborate these findings and to examine whether the effects of music therapy are enduring.

PLAIN LANGUAGE SUMMARY

Music therapy for autistic people

Review question

We reviewed the evidence about the effect of music therapy for autistic people. We compared results from people receiving music therapy (or music therapy added to standard care) with results from people receiving a similar therapy without music ('placebo' therapy), standard care or no therapy at all.

Background

Autism is a lifelong neurodevelopmental condition that affects how people perceive the world around them, and how they communicate with and relate to others. Thus, social interaction and social communication are among the central areas of difficulty for autistic people. Music therapy uses music experiences and the relationships that develop through them to enable people to relate to others, to communicate, and to share their feelings. In this way, music therapy addresses some of the core problems of autistic people. Music therapy has been applied in autism since the early 1950s. Its availability to autistic people varies across countries and settings. The application of music therapy requires specialised academic and clinical training. This helps therapists in tailoring the intervention to the specific needs of the person. We wanted to investigate whether music therapy helps autistic people compared with other options.

Search date

The evidence is current to August 2021.

Study characteristics

We included 16 new studies in this update, so the evidence in this review now rests on 26 studies with a total number of 1165 participants. The studies examined the short- and medium-term effect of music therapy interventions (three days to eight months) for autistic children, youth, and young adults in one-to-one or group settings. None of the studies reported funding by an agency with a commercial interest in the result of the studies; reported sources of support included governmental, university and foundation funding; in three studies, support was provided by a music therapy association.

Key results

Music therapy compared with 'placebo' therapy or standard care probably increases the chance of overall improvement by the end of therapy, likely improves quality of life and total autism symptom severity immediately after therapy, and probably does not increase adverse events. From the available evidence, we cannot tell whether music therapy has any effects on social interaction, and verbal and non-verbal communication at the end of therapy.

Quality of the evidence

The evidence we found in this review is of very low to moderate certainty. This means that future research may change these findings and our confidence in them. We found that music therapy is probably effective regarding global improvement, quality of life, total autism symptom severity and adverse events measured at the end of therapy based on the moderate certainty of the evidence in these domains. It remains unclear whether music therapy has an effect on social interaction, non-verbal communication and verbal communication at the end of therapy since the certainty of evidence was low to very low. Reasons for limited certainty of the evidence were issues with study design and blinding (i.e. those who applied outcome measures often knew whether or not participants had received music therapy, which may have influenced their assessments).

Authors' conclusions

Music therapy compared with 'placebo' therapy or standard care probably increases the chance of overall improvement by the end of therapy. It also probably helps to enhance quality of life, and lessen symptom severity. Music therapy probably does not increase adverse events. We cannot tell whether music therapy may help with social interaction, non-verbal communication and verbal communication at the end of therapy. Most of the included studies featured interventions that correspond well with music therapy in clinical practice concerning methods and settings. This new evidence is important for autistic people and their families as well as for policymakers, service providers and clinicians, to help in decisions around what intervention to choose, and in the planning of resources. More research with adequate design (i.e. producing reliable evidence) looking at areas that matter to autistic people is needed. Because long-term outcomes of therapy matter to autistic people and their families, it is important to specifically examine how long the effects of music therapy last.

Summary of findings 1. Music therapy compared with placebo therapy or standard care for autistic people

Music therapy compared with placebo therapy or standard care for autistic people

Population: individuals with a diagnosis of autism spectrum disorder
Settings: outpatient therapy centre, hospital, school, summer camp or home; individual and group setting
Intervention: music therapy
Comparison: placebo therapy or standard care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk	Music therapy versus placebo therapy or standard care			
	Risk with placebo or standard care	Risk with music therapy				
Global improvement Follow-up: immediately post-intervention (M = 3.4 months, SD = 2.4)	**Low-risk population**[a]		**RR 1.22** (1.06 to 1.40)	583 (8 studies)	⊕⊕⊕⊖ **Moderate**[b]	Higher scores represent greater improvement.
	430 per 1000	**525 per 1000** (456 to 602)				
	High-risk population[a]					
	800 per 1000	**976 per 1000** (848 to 1000)				
Social interaction Follow-up: immediately post-intervention (M = 3.5 months, SD = 2.4)	-	The mean social interaction score at immediately post-intervention in the intervention groups was **0.26 standard deviations higher** (0.05 lower to 0.57 higher)	-	603 (12 studies)	⊕⊕⊖⊖ **Low**[c]	Higher scores represent higher social interaction capabilities.

						Small to medium effect size according to Cohen 1988
Non-verbal communication Follow-up: immediately post-intervention (M = 4.2 months, SD = 2.4)	-	The mean non-verbal communication score at immediately post-intervention in the intervention groups was **0.26 standard deviations higher** (0.03 lower to 0.55 higher)	-	192 (7 studies)	⊕⊕⊖⊖ **Low**[d]	Higher scores represent higher non-verbal communication capabilities. Small to medium effect size according to Cohen 1988
Verbal communication Follow-up: immediately post-intervention (M = 3.2 months, SD = 2.8)	-	The mean verbal communication score at immediately post-intervention in the intervention groups was **0.30 standard deviations higher** (0.18 lower to 0.78 higher)	-	276 (8 studies)	⊕⊖⊖⊖ **Very low**[e]	Higher scores represent higher verbal communication capabilities. Small to medium effect size according to Cohen 1988
Quality of life Follow-up: immediately post-intervention (M = 3.3 months, SD = 1.5)	-	The mean quality of life score at immediately post-intervention- in the intervention groups was **0.28 standard deviations higher** (0.06 to 0.49 higher)	-	340 (3 studies)	⊕⊕⊕⊖ **Moderate**[f]	Higher scores represent higher quality of life. Small to medium effect size according to Cohen 1988
Total autism symptom severity	-	The mean total autism symptom severity score at immediately post-intervention in the intervention	-	575 (9 studies)	⊕⊕⊕⊖ **Moderate**[b]	Higher scores represent higher symptom severity.

Follow-up: immediately post-intervention (M = 3.6 months, SD = 2.1)		groups was **0.83 standard deviations lower** (1.41 to 0.24 lower)				Large effect size according to Cochrane 1988
Adverse events Any serious or non-serious adverse event Follow-up: immediately post-intervention (M = 4.0 months, SD = 1.4)	**Low-risk population**[a]		**RR 1.52** (0.39 to 5.94)	326 (2 studies)	⊕⊕⊕⊖ **Moderate**[f]	Higher scores represent higher numbers of adverse events. Adverse events reported are hospitalisation periods, typically planned and short-term. One study with 36 participants reported no adverse events and was not included in the RR analysis.
	0 per 1000	**0 per 1000** (0 to 0)				
	High-risk population[a]					
	24 per 1000	**37 per 1000** (9 to 150)				

*The basis for the **assumed risk** is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the intervention group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **M:** Mean; **RR:** Risk ratio; **SD:** Standard deviation.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of the effect.

[a]Typical risks are not known, so we chose the risk from included studies providing the second highest (Kim 2008) for a high-risk population and the second lowest (Porter 2017) for a low-risk population for the outcome 'Global improvement' (Schünemann 2021). For the outcome of 'Adverse events', where only two studies were included, we based the risk of the high-risk population on Bieleninik 2017 and that of the low-risk population on Porter 2017.
[b]We downgraded the certainty of the evidence by one level for risk of bias (limitations in the designs such as poorly reported randomisation, blinding of outcomes, incomplete outcome data).
[c]We downgraded the certainty of the evidence by one level for risk of bias and one level for imprecision (wide CI: 95% CI included no effect and the upper confidence limit crossed an effect size of 0.5; GRADEpro GDT).

[d]We downgraded the certainty of the evidence by two levels for imprecision (wide CIs) and because the total number of participants in this outcome was lower than 400.
[e]We downgraded the certainty of the evidence by one level for risk of bias and two levels for imprecision (wide CIs), and because the total number of participants in this outcome was lower than 400.
[f]We downgraded the certainty of the evidence by one level for imprecision because the total number of participants in this outcome was lower than 400.

BACKGROUND

Description of the condition

Autism is a complex neurodevelopmental condition that usually manifests in early childhood and persists throughout life. When following a medical paradigm, and according to the criteria of the International Classification of Diseases and Related Health Problems, 11th edition (ICD-11) (WHO 2021), and the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) (APA 2013), autism spectrum disorder (ASD) is characterised by 'persistent deficits in social communication and social interaction across multiple contexts', and by the presence of 'restricted, repetitive patterns of behavior, interests, or activities'. For a diagnosis of ASD, children must show symptoms of ASD since early childhood (i.e. before the age of three) (APA 2013; WHO 2021). In some instances, these symptoms may only be detectable later when social demands become intractable, or may continue to be masked through learned strategies (APA 2013) in an attempt to mimic neurotypical behaviours.

The prevalence of ASD has considerably risen over the last decades. While the first epidemiological study estimated a prevalence of the condition as lower than 0.5% in young children (Lotter 1966), the latest estimates of the Centers for Disease Control and Prevention reported that one in 54 children in the United States may be on the autism spectrum (Maenner 2020). The increased prevalence rates are attributable to the broadening of the diagnostic criteria, diagnostic switching from other developmental disabilities, service availability, and awareness of the condition among the community and professionals (Elsabbagh 2012; Lyall 2017). Of note, ASD is more commonly diagnosed among males than females, with a ratio of 4:1 (Maenner 2020).

The clinical picture is widely variable in presentation, severity, and hence levels of support needed. Additionally, ASD may be accompanied by co-occurring conditions, such as intellectual disability (ID), language impairments, as well as other neurodevelopmental, mental, and behavioural disorders (APA 2013). The most frequent co-occurring mental health conditions are attention-deficit hyperactivity disorders (ADHD), anxiety disorders, sleep-wake disorders, depression, disruptive, impulse-control, and conduct disorders (Lai 2019). Autistic people might be more vulnerable to negative life experiences (Griffiths 2019) and to the development of post-traumatic stress symptoms (Rumball 2020). As a consequence, outcome domains beyond the core symptoms of ASD, such as depression, anxiety, or quality of life, are increasingly receiving more attention in autism research.

As autistic ways of communicating and being social deviate from neurotypical socialising, approaches following a medical model tend to seek to change this deviation. Within such a model, challenges emerging from being autistic are situated within the autistic individuals rather than the environment, culture or society surrounding them. The medical model has been criticised by scholars as well as by the autistic community (De Jaegher 2013; Milton 2012). Instead, a social or cultural model of autism has been suggested (Sinclair 2010; Sinclair 2012). A social model of understanding autism looks at autistic characteristics as part of human diversity and understands social interaction as a shared responsibility and participatory practice (De Jaegher 2007). Hence, challenges causing dysfunction in social interaction can also be located outside the autistic person and might require changes from the environment rather than solely from the individual. Accordingly, the enabling and disabling impact any given interaction, context or society can hold for autistic people needs to be considered when defining autism or interacting with autistic people (Milton 2019).

Regarding the terminology used in autism research, there is an ongoing debate on the type of language that is most appropriate and most respectful to people with a diagnosis of ASD, their families and caregivers. A growing body of literature documents that person-first language (e.g. 'people with ASD') may actually increase effects of stigmatisation for autistic people (Bottema-Beutel 2021; Gernsbacher 2017), and that people with a diagnosis of ASD themselves often prefer using identity-first language (e.g. "autistic individuals") as a means of showing that autism is a central part of their identity rather than something that needs to be fixed or cured (Bury 2020; Kenny 2016). This preference has also been expressed by autistic people and their families who have been consulted by the authors while conducting this review. Considering these contexts and perspectives, we chose to use identity-first language in this review.

Depending on the way autism is conceptualised - either as a set of cognitive or behavioural deficits (APA 2013) or as a social construction and as a description of a culturally filtered experience (Milton 2019) - therapeutic aims and approaches will differ. Following a medical paradigm, psychosocial and behavioural therapies are considered the first-line evidence-based treatments for people with a diagnosis of ASD. These therapy approaches traditionally aim at achieving changes regarding the way autistic people communicate and interact with others and often follow a normalising agenda which tries to lessen or remove outward signs of autism. Contrary to this, some of the same therapies may follow a maximising agenda (Winter 2012) where the aims of any intervention are to maximise an individual's capabilities as an autistic person.

A variety of music therapy approaches have been developed for working with autistic people, many of them defined as relational or child-led (Carpente 2009; Geretsegger 2015; Schumacher 1994), following the individual's strengths and resources and allowing for participatory processes in the development of social interaction and understanding, thus more aligned with a maximising agenda and a cultural or social model of autism.

As symptom change is often assessed as a primary outcome in scientific research, a normalisation agenda might be considered to form the epistemological background of music therapy research as well. However, music therapy research also combines these two agendas (Pickard 2020) by applying music therapy approaches striving for maximisation, while concurrently using outcomes measuring neurotypical social behaviour and communication and general domains such as quality of life (see, for example, Bieleninik 2017). Thus, music therapy can be seen as relating to the coexisting "dual nature of autism" (Lai 2020), being categorised as medical condition leading to developmental disability and at the same time being an example of neurodivergent development forming identity and culture.

Description of the intervention

Music therapy for autistic people is often provided as individual therapy, although there are also reports of group-based and peer-

mediated interventions (e.g. Boso 2007; Ghasemtabar 2015; Kern 2006; Kern 2007; LaGasse 2014; Mateos-Moreno 2013). Family-centred approaches, where parents or other family members are included in therapy sessions (Oldfield 2012; Pasiali 2004; Thompson 2014; Thompson 2012) or trained in relevant music-based techniques for social engagement (Gottfried 2016), have increasingly become an important part of music therapy for autistic children, especially to help generalise skills acquired in therapy to everyday contexts, that is, to transfer these skills from the therapy context to new and different settings.

Music therapy has been defined as "a systematic process of intervention wherein the therapist helps the client to promote health, using music experiences and the relationships that develop through them as dynamic forces of change" (Bruscia 1998, p. 20). Music therapy approaches for autistic people are based on sensory-perceptional, developmental, creative, behavioural, and educational conceptualisations (Bergmann 2016). Accordingly, aims in music therapy are wide including the work on communication and interaction, sensory processing and integration, affect regulation, identity formation as well as creative and recreational needs that can lead to an increased quality of life. Active music-making with a variety of instruments that are easy to play is widely used, involving the client and the therapist in joint musical play. Central music therapy techniques include free and structured improvisation, recreating songs and vocalisation, or songwriting. Listening to pre-recorded or live music played by the therapist can be used for e.g. relaxation purposes or, in the context of behavioural approaches of music therapy, focusing on training of specific skills. Some music therapy approaches also include movement activities or story-telling. The delivery of music therapy varies in its degree of structuredness: while behavioural approaches often make use of fixed manuals specifying training phases and materials (e.g. Lim 2011), developmental or improvisational approaches are usually less pre-structured. However, there are also some flexible yet systematic treatment guidelines for improvisational music therapy in autism which specify core therapeutic principles and techniques (Geretsegger 2015; Kim 2006; Thompson 2014; Wigram 2006).

Music therapy has been applied in autism since the early 1950s (Fusar-Poli in press; Reschke-Hernández 2011), but its availability to autistic individuals varies across countries, depending on other factors such as age or educational setting (Kern 2017). The application of music therapy requires specialised academic and clinical training, typically achieved through Bachelor and Master's level degree courses in music therapy which usually lead to accreditation with professional associations or governmental registries, or both. Training courses in music therapy not only teach clinical music therapy techniques, but also aim at developing the therapist's personality and clinical sensitivity, which is necessary to apply music therapy responsibly. Thus, this specialised training enables music therapists to tailor their methods and techniques to meet individual therapeutic goals and needs (Fusar-Poli in press).

How the intervention might work

The processes that occur within musical interaction may help autistic people to develop communication skills and the capacity for social interaction. Through engaging in musical interaction, participants in music therapy can shift between verbal, non-verbal and pre-verbal modes of communication. Thus, musical interaction can be understood and described as a means for verbal people to access sensory experiences and for people without spoken language to interact communicatively without words. It enables all to engage on a more emotional, relationship-oriented level than may be accessible through verbal language (Alvin 1991). Behaviouristic and educational approaches typically use music activities to motivate the child and to reinforce targeted behaviour. Developmental approaches often use music to focus on the sensory, motor-coordination and affective aspects of music-making, e.g. through intra- and inter-personal synchronisation experiences (Berger 2002; Schumacher 2019). In improvisational approaches, therapists attune to the child's intrinsic way of sound-making and moving, using the shared history of musical interaction and jointly developed musical activities to motivate and engage the child in interactive processes (Geretsegger 2015; Holck 2004). Listening to music within music therapy also involves an interactive process that often includes selecting music that is meaningful for the person (e.g. relating to an issue that the person is occupied with) and, where possible, reflecting on personal issues related to the music or associations brought up by the music. For those with verbal abilities, verbal reflection on the musical processes is often an important part of music therapy (Wigram 2002).

There are several psychological theories and neurobiological models that aim to explain the mechanisms through which music therapy helps autistic individuals (Fusar-Poli in press). One area of research underpinning the potential of music therapy in autism is based on findings suggesting that motor timing and sensorimotor integration are disrupted in autistic people (De Jaegher 2013; Sharda 2018), which may contribute to broader challenges in interacting with others (Mössler 2019). Functional neuroimaging studies with autistic individuals showed an overconnectivity between sensory brain networks which is related to the sensory processing differences and multisensory integration difficulties (Chen 2020). Thus, sensorimotor integration facilitated by musical interaction may lead to modulation of atypical sensory processing, which may in turn enhance social communication (Thye 2018).
Another, related rationale for the use of music therapy for individuals with communication disorders is based on the findings of infancy researchers such as Stern and Trevarthen who describe sound dialogues between mothers and infants using 'musical' terms (Stern 1985; Stern 1989; Stern 2010; Trevarthen 1999b). When describing tonal qualities, researchers use the terms pitch, timbre, and tonal movement and, when describing temporal qualities, they speak of pulse, tempo, rhythm, and timing (Wigram 2002). Trevarthen 1999a describes the sensitivity of very young infants to the rhythmic and melodic dimensions of maternal speech, and to its emotional tone, as demonstrating that we are born ready to engage with the 'communicative musicality' of conversation. The experience of attunement through synchronisation in timing, tonality or affective dynamics shapes the attachment between infant and caregivers and has been suggested as influencing the development of social understanding (Greenspan 2007; Stern 1985; Trevarthen 2011). These premises allow music to act as an effective medium for engaging in non-verbal social exchange for autistic children and adults. Communicative behaviours, such as joint attention, eye contact, and turn-taking, are characteristic events in shared, active music-making and, therefore, inherent components of music therapy processes. Recent research has shown that musical and emotional attunement within music therapy processes can support social responsiveness in autistic children (Mössler 2019; Mössler 2020). In addition to music's potential to stimulate communication (as

described for vocal communication in Salomon-Gimmon 2019), music therapists use music, especially improvisational music-making, to provide autistic people with opportunities to experience structure combined with measured flexibility, thus helping them to find ways of coping in less predictable situations that will typically pose challenges for them (Wigram 2009).

The potential for predictability and anticipation brought about by musical structures is an element also used in behavioural approaches where music is utilised as a stimulus facilitating the perception and production of speech and language and enhancing communication skills (Lim 2010; Lim 2011). Another rationale for using music in this way is the increased attention and enjoyment observed in autistic individuals when presented with musical as opposed to verbal stimuli (Buday 1995; Lim 2010; Lim 2011).

Why it is important to do this review

This is an update of a Cochrane review first published in 2006 (Gold 2006) and previously updated in 2014 (Geretsegger 2014). The first version of this review concluded that music therapy may help autistic children to improve their communicative skills, but also noted that more research was needed to investigate the effects of music therapy in typical clinical practice and within longer periods of observation (Gold 2006). In the 2014 update of this review, we found that music therapy may help autistic children to improve their skills in social interaction, verbal communication, initiating behaviour, and social-emotional reciprocity; we also concluded that more research with larger samples addressing relevant outcomes through standardised scales was needed to corroborate these findings and to examine long-term effects of music therapy as well as effects of music therapy for adolescents and adults (Geretsegger 2014).

More recently, further systematic reviews have appeared, often with limited scope (e.g. Shi 2016 focusing on only Chinese data), methodological flaws (e.g. Whipple 2012 where designs of included studies lacked homogeneity and included sample sizes of only one), or providing only narrative summaries (e.g. De Vries 2015), thus highlighting the continued need for an updated, comprehensive review. Furthermore, considerable changes have occurred in the knowledge about ASD in recent years, and a number of new studies of music therapy for autism were published since the 2014 version of this review, which necessitated an update of the previous review. We conducted the current update to summarise and evaluate these new studies in order to provide comprehensive and up-to-date conclusions, as well as implications for practice and research that are based on the most recent findings. This information is highly relevant for autistic individuals and their families as well as for policymakers, service providers and clinicians, to help in decisions around the types and amount of intervention and support that should be provided, and in the planning of resources.

OBJECTIVES

To review the effects of music therapy, or music therapy added to standard care, for autistic people.

METHODS

Criteria for considering studies for this review

Types of studies

All relevant randomised controlled trials (RCTs), quasi-randomised trials and controlled clinical trials (CCTs), including cluster-trials were considered for inclusion. Studies using single-case experimental designs were included if they also met the definition of RCTs or CCTs, that is, if the different interventions were provided in a different order to different participants (i.e. cross-over RCTs/CCTs). Studies in which all participants received interventions in the same order (i.e. case series) were excluded.

Types of participants

Individuals of any age who were diagnosed with ASD as defined in DSM-5 (APA 2013) or ICD-11 (WHO 2021) criteria, whether identified by a psychological assessment or a psychiatric diagnosis, were considered for inclusion. Moreover, we included individuals diagnosed with pervasive developmental disorders, as defined in ICD-10 criteria (WHO 1994) or in previous versions of the DSM, including childhood autism, atypical autism, Asperger's syndrome, and pervasive developmental disorder not otherwise specified, as these previous diagnostic labels are now included in the category of ASD in DSM-5 and ICD-11. Individuals with Rett's disorder or childhood disintegrative disorder were not included as they have been excluded from the ASD diagnostic category in the current classifications, given their significantly different clinical course.

Types of interventions

Interventions included music therapy (i.e. regular sessions of music therapy involving music experiences and relationships developing through them as defined above, delivered by a professional music therapist).

Comparators

Interventions were compared with either 'placebo' therapy (i.e. a similar intervention without the elements specific to music therapy, e.g. play therapy without music, or music listening without interaction with a music therapist; the concept of attention placebo in psychotherapy research is discussed in Kendall 2004), no treatment, or standard care control; or music therapy added to standard care compared with standard care (with or without 'placebo' therapy).

Types of outcome measures

To ensure that all user-important outcomes were addressed (McKenzie 2021), and to update our approach in correspondence with changes that occurred in the knowledge and nosological classification of the condition in recent years (see Differences between protocol and review), we adapted the outcome categories used in the previous version of the review, as described below. In our adaptations, we also sought to broaden outcome areas in order to not only address specific skills (e.g. social adaptation; communicative skills such as eye contact, imitating gestures or words), but also wider areas of capacity (e.g. adaptive behaviour in more than just the social domain; communication including all domains of verbal or non-verbal communication, pragmatics, language structure, and communication behaviours such as withdrawal within a group).

We considered the broad-based measures 'global improvement' (binary) and 'total autism symptom severity' (continuous) as primary outcomes. Although not endorsed when applying a social-model approach to autism, measures relating to these overall categories are still considered important in a medical-model perspective on autism which is likely to be relevant for many policymakers, service providers and clinicians. As in the 2014 version of this review, we also regarded outcome measures in all areas of social communication as primary outcomes as they refer to the core characteristics defining ASD (social interaction, non-verbal communication, verbal communication). In addition, we moved the category of 'quality of life' (secondary outcome 'quality of life in school, home, and other environments' in the 2014 version of this review) into primary outcomes due to the increased relevance of this outcome to autistic individuals and their families, as demonstrated in recent studies and reports (e.g. McConachie 2015; Provenzani 2020). To keep this review focused and manageable for users (McKenzie 2021), we merged previously separate outcomes that concern specific sub-skills of social interaction ('initiating behaviour', 'social-emotional reciprocity', 'joy'), with the wider category of 'social interaction'. Finally, we retained 'adverse events' as a primary outcome category.

We regarded other commonly examined outcome measures in areas not specific to defining ASD characteristics as secondary outcomes. The outcome 'social adaptation skills' was re-labelled as 'adaptive behaviour'. In order to address outcomes that are regarded as highly relevant by autistic people, their family members and professionals (Lipinski 2019; McConachie 2015) and that were evaluated in included studies, we newly added 'identity formation' (including self-esteem) and 'depression' as secondary outcome categories.

Finally, we removed the outcome category 'hyperacusis (hypersensitivity to sound)', as we did not find it measured in any study, or mentioned in any review.

Data sources could have included non-standardised or standardised instruments (for a review of relevant standardised instruments, see Ozonoff 2005; McConachie 2015; Provenzani 2020), parent or teacher report, or school records. Data from rating scales were only included if the instrument was either a self-report or completed by an independent rater or relative (i.e. not the therapist, unless reconfirmed by an independent rater).

Primary outcomes

Primary outcomes included the following.

1. Global improvement: binary (improved versus not improved or unknown, on a scale measuring clinical global impressions or on a global measure used as primary outcome in a study);
2. Social interaction: continuous;
3. Non-verbal communication: continuous;
4. Verbal communication: continuous;
5. Quality of life: continuous; could be measured in various contexts (school, home, other) and with varying scope (individual, family);
6. Total autism symptom severity: continuous;
7. Adverse events: binary (any adverse event/no adverse event), as defined by study authors.

Secondary outcomes

Secondary outcomes included the following.

8. Adaptive behaviour: continuous; this could be measured as positive adaptive behaviours (enabling a person to get along in their environment with greatest success and least conflict with others) or as maladaptive, dysfunctional behaviours (which stop a person from adapting to new or difficult circumstances, including 'restricted and repetitive behaviours');
9. Quality of family relationships: continuous;
10. Identity formation: continuous; including self-esteem and related concepts;
11. Depression: continuous;
12. Cognitive ability: continuous; including attention and concentration.

Changes in generalised skills that are measured outside of the immediate therapy context pose the biggest challenge for any interventions for autism (Warren 2011). Generalised outcomes refer to changes that generalise to other behaviours and to other contexts across settings, people, or materials. In the Summary of findings 1, we report the results of seven generalised outcomes (all listed under Primary outcomes) measured immediately post-intervention.

We grouped outcome time points as follows: during the intervention (previously labelled "within sessions/non-generalised"); immediately post-intervention; one to five months post-intervention; six to 11 months post-intervention; 12 to 23 months post-intervention; and 24 to 35 months post-intervention. Where outcomes were measured at multiple time points during the course of therapy, we used mean values of all data from the second therapy session onwards.

Search methods for identification of studies

We ran the searches for this update in July 2020 and again in August 2021. We revised the original search strategy by removing redundant search terms, and by adding relevant database sources which were either not available at the time of search for the previous update (e.g. MEDLINE Epub Ahead of Print) or not routinely included previously (e.g. trial registers). Where possible, searches were limited to the period since the last update (2013 onwards). For newly added databases, searches were conducted since their inception.

Electronic searches

The Cochrane Developmental, Psychosocial and Learning Problems Group Information Specialist, Margaret Anderson, conducted systematic searches in the following databases for randomised controlled trials and controlled clinical trials without language or publication status restrictions.

1. Cochrane Central Register of Controlled Trials (CENTRAL; 2021, Issue 8), part of the Cochrane Library (searched 4 August 2021).
2. MEDLINE Ovid (1946 to July week 4 2021).
3. MEDLINE In-Process & Other Non-indexed Ciations Ovid (1946 to 3 August 2021).
4. MEDLINE EPub Ahead of Print Ovid (3 August 2021).
5. Embase Ovid (1974 to 3 August 2021).
6. LILACS (lilacs.bvsalud.org/en/; searched 4 August 2021).

7. APA PsycINFO Ovid (1806 to July week 4 2021).
8. CINAHL EBSCOhost (1937 to 4 August 2021).
9. ERIC EBSCOhost (1966 to 4 August 2021).
10. Sociological Abstracts Proquest (1952 to 4 August 2021).
11. Proquest Global Dissertations & Theses (searched 4 August 2021).
12. Proquest Music Periodicals Database (1996 to 4 August 2021).
13. Proquest Performing Arts Periodicals Database (1864 to 4 August 2021).
14. RILM Abstracts of Music Literature Online (1967 to 4 August 2021).
15. *Cochrane Database of Systematic Reviews* (CDSR; 2021, Issue 8), part of the Cochrane Library (searched 4 August 2021).
16. Epistemonikos (www.epistemonikos.org/en/; searched 4 August 2021).
17. ClinicalTrials.gov (clinicaltrials.gov/ct2/home; searched 5 August 2021).
18. WHO International Clinical Trials Registry Platform (apps.who.int/trialsearch/; searched 5 August 2021).

Detailed search strategies for this update are reported in Appendix 1. Details of the previous search strategies are available in Geretsegger 2014.

Searching other resources

Adverse events

We did not perform a separate search for adverse events. We considered adverse events described in included studies only.

Searching reference lists

We checked the bibliographies of included studies and relevant reviews (Accordino 2007; Ball 2004; Brondino 2015; De Vries 2015; Pater 2017; Reschke-Hernández 2011; Shi 2016; Simpson 2011; Weitlauf 2014; Whipple 2004; Whipple 2012) for further references to relevant trials.

Searching by contacting individuals or organisations

We contacted experts and organisations in the field through correspondence in researcher networks, conferences and social media to gather information on ongoing trials and any relevant material not captured by our searches. Where necessary, we contacted authors of key papers and abstracts to request further information about their trials.

Data collection and analysis

Selection of studies

We used Cochrane's Screen4Me workflow to help assess the search results. Screen4Me comprises three components: known assessments — a service that matches records in the search results to records that have already been screened in Cochrane Crowd and been labelled as an RCT or as Not an RCT; the RCT classifier — a machine learning model that distinguishes RCTs from non-RCTs and, if appropriate, Cochrane Crowd — Cochrane's citizen science platform where the Crowd help to identify and describe health evidence. For more information about Screen4Me and the evaluations that have been done, please go to the Screen4Me webpage on the Cochrane Information Specialist's portal: https://community.cochrane.org/organizational-info/resources/resourcesgroups/ information-specialists-portal. In addition, more detailed information regarding evaluations of the Screen4Me components can be found in the following publications: Marshall 2018; Noel-Storr 2020; Noel-Storr 2021; Thomas 2020.

Four authors (CE, LFP, MG, GV) independently inspected all titles and abstracts identified from the search in such a way that each record was screened by two authors. We obtained potentially relevant papers and resolved any disagreement about eligibility through discussion or consultation with the other authors. For non-English study reports, we provided for their translation. We recorded the reasons for excluding trials.

We recorded the selection process in sufficient detail to produce a PRISMA flow diagram (Liberati 2009).

Data extraction and management

Four reviewers (CE, LFP, MG, GV) independently performed data extraction using a data collection form so that data from each study were extracted by two reviewers. We made sure that studies in which any of the reviewers were involved with were dealt with by two other reviewers not involved in these studies. The data collection form was initially piloted to ensure feasibility and included details on study design, participants, interventions, outcomes including measurement time points and allocation to outcome categories, and funding sources. Any disagreements were resolved by discussion, or consultation with the other reviewers, or both. When necessary, we contacted the study authors to provide missing data.

Assessment of risk of bias in included studies

Four authors (KM, MG, LFP, GV) independently assessed the risk of bias of each included study using the Cochrane risk of bias tool (Higgins 2011). We made sure that studies where any of the reviewers were involved were dealt with by other reviewers not involved in these studies. Any disagreements were resolved by discussion, or consultation with the other reviewers, or both.

For each included study, we presented the risk of bias assessments in a table where the judgement of the review authors (low, high or unclear risk of bias) was followed by a text box providing details on the available information that led to each judgement.

We assessed the following items:

1. random sequence generation;
2. allocation concealment;
3. blinding of participants and personnel;
4. blinding of outcome assessment;
5. completeness of outcome data;
6. selective reporting; and
7. other sources of bias.

The criteria for assigning judgements of high, low and unclear risk of bias are provided in Appendix 2.

Measures of treatment effect

Where available, we used individual participant data (IPD) in order to calculate measures of treatment effect consistently.

Binary data

We calculated the risk ratio (RR) and corresponding 95% confidence interval (95% CI) for binary outcomes. The number needed to treat for one beneficial outcome was calculated, where appropriate.

Continuous data

We preferred endpoints over change scores. If IPD were available, the distributions of values were visually checked for skewness. Where skewness was found, we attempted to remove it by log-transformation. We then examined how log-transformation influenced the effect size estimate and used the more conservative estimate.

We calculated the standardised mean difference (SMD) and corresponding 95% CI for all continuous outcomes. When combining different scales for the same outcome, it was necessary to standardise the effects in order to make them comparable. When combining results for the same scale, either the mean difference (MD) or SMD could have been used. We decided to use SMD in order to facilitate the interpretation of effect sizes as small (up to 0.2), medium (around 0.5) or large (0.8 and above) based on guidelines that are commonly used in the behavioural sciences (Cohen 1988; Schünemann 2021a). In the absence of any anchor-based minimally clinical important differences (MCIDs) for the outcomes in this review, the general guidelines for the behavioural sciences developed by Cohen 1988 state that an effect size needs to reach at least a level of 0.2 to be regarded as potentially meaningful; effects smaller than 0.2 may be negligible. It is noted that the choice of SMD or MD does not usually affect the significance level of the results and the authors cautiously assessed whether this was the case.

All SMDs, regardless of whether the study was a parallel or a cross-over design, were standardised by the pooled standard deviation between participants, rather than the standard deviation of the difference within participants. This is the standard procedure, which enables comparisons of different scales and facilitates interpretation of the magnitude of effects (Cohen 1988; Gold 2004). The calculation of the standard error then depends on the study design. For parallel designs, the standard error was calculated using the standard formulae for SMDs as implemented in RevMan Web (RevMan Web 2020). For cross-over studies, outcomes are usually positively correlated within participants; we assumed a correlation of 0 as a conservative estimate; this avoided giving too high weight to small studies and also enabled use of standard methods for SMDs within RevMan (Elbourne 2002; Higgins 2021).

For studies where outcomes were measured on several occasions during each treatment intervention, we used the mean of all measurements from the second occasion onwards. Where the same outcome was measured on multiple occasions using the same scale, we calculated the mean and the pooled SD and entered these into RevMan. Where the same outcome was measured on multiple occasions using different scales, we calculated a mean effect size for that study outcome and entered that into RevMan (along with SD 1 and mean in control 0).

In comparison to the previous review update, these procedures ensure better consistency and transparency, but also tend to show more conservative results. Thus, a study that in Geretsegger 2014 showed a significant effect (using the generic inverse variance method in RevMan and possibly a change score or a positive correlation estimate) might show no effect in this update.

Unit of analysis issues

Cluster-RCTs

For cluster-RCTs, we adjusted the sample size according to the design effect, based on an intraclass coefficient calculated from IPD, if available.

Cross-over trials

The appropriateness of cross-over designs is difficult to assess. In general, autism as a lifelong condition lends itself well to such designs. However, it is less clear how lasting any effects of music therapy may be. In general, we judged cross-over designs as appropriate unless there was clear evidence to the contrary (e.g. a clearly irreversible outcome). We therefore combined the results of cross-over trials with the results of parallel-group trials, and used data from all periods in order to retain a maximum of information provided by those studies. Data from washout periods in cross-over studies were excluded from the analysis.

Multiple treatment arms

For studies including more than one relevant music therapy group or more than one relevant control group, we combined the data of the relevant groups by calculating a weighted mean and pooled SD.

Dealing with missing data

We assessed loss to follow-up and dropouts in the included studies as reported in the risk of bias tables. Where unclear, we contacted the study authors to confirm any loss to follow-up and dropouts in their studies. We applied an intention-to-treat analysis (except for adverse events) for available cases and did not impute missing values for continuous outcomes. We are aware that this may introduce bias if being lost to follow-up is related to a participants' response to intervention (Moher 2010). Therefore, we examined the impact of studies with high risk of bias due to dropout using sensitivity analyses, where these studies were excluded.

Assessment of heterogeneity

Because statistical tests of heterogeneity have low power, particularly when the number of studies is low, we relied primarily on descriptive analyses of heterogeneity. We visually inspected forest plots for consistency of results and calculated the I^2 statistic (Higgins 2002), which describes the proportion of variation in point estimates that is due to heterogeneity rather than sampling error, and followed suggested threshold bands for interpreting the I^2 statistic which define 0% to 40% as "might not be important"; 30% to 60% as "may represent moderate heterogeneity"; 50% to 90% as "may represent substantial heterogeneity"; and 75% to 100% as "considerable heterogeneity" (Deeks 2021). We supplemented this by calculating the Chi^2 statistic to determine the strength of evidence that the heterogeneity was genuine. We investigated possible sources of heterogeneity when it was detected.

Assessment of reporting biases

We used funnel plots to investigate any relationship between effect size and study precision in cases where 10 or more studies were pooled for an outcome. With other design aspects equal, a funnel

plot would be symmetric within chance variation in the absence of publication bias; a noticeable asymmetry may therefore indicate a strong publication bias. However, because the method may not work well when larger studies differ in other design aspects, as well as because of its subjective interpretation, we did not interpret a lack of an apparent asymmetry as evidence of absence of publication bias.

Data synthesis

Using RevMan Web (RevMan Web 2020), we conducted meta-analyses utilising RRs for dichotomous outcomes and SMDs for continuous outcomes. A fixed-effect model was initially used for all analyses. If a common effect size was not tenable because a substantial amount of heterogeneity (i.e. 50% or higher; Deeks 2021) was identified that could not be explained by clinical subgroups in the outcome domain immediately post-intervention (see Subgroup analysis and investigation of heterogeneity), we chose a random-effects model. Where we conducted fixed-effect analyses, we also examined whether random-effects analyses would have altered the results by conducting sensitivity analyses, and reported any differences in the Effects of interventions section. We used the inverse variance method, which is most commonly used, in random-effects analyses of dichotomous outcomes and in all analyses of continuous outcomes. In fixed-effect meta-analyses of dichotomous outcomes, we used the Mantel-Haenszel method, which is the default method in RevMan Web and is commonly preferred because it has better statistical properties when there are few events (Deeks 2021).

Subgroup analysis and investigation of heterogeneity

When substantial heterogeneity was identified ($I^2 \geq 50\%$), we examined the impact of clients' age (children versus adolescents or adults), intensity of therapy (i.e. number and frequency of music therapy sessions), and treatment quality (i.e. adequate music therapy methods; adequate training of therapists; see definitions specified in Appendix 2, 'Other bias') in subgroup analyses.

Sensitivity analysis

We conducted sensitivity analyses to determine the impact of attrition bias risk by removing studies at high risk of attrition bias. We also investigated the impact of the choice of model by conducting a random-effects analysis where fixed analysis was chosen and comparing the findings.

Summary of findings and assessment of the certainty of the evidence

We created a summary of findings table for our main comparison: music therapy compared with placebo therapy or standard care. We included the following primary outcomes, assessed immediately post-intervention: global improvement; social interaction; non-verbal communication; verbal communication; quality of life; total autism symptom severity; adverse events.

Four review authors (KM, MG, LFP, GV) assessed the overall certainty of the body of evidence using the GRADE approach (Schünemann 2013). We made sure that studies in which any of the reviewers were involved with were dealt with by two other reviewers not involved in these studies. Any disagreements were resolved by discussion, or consultation with the other review authors, or both. The certainty of the evidence for each outcome was graded as high, moderate, low, or very low, according to the presence of the following five criteria: risk of bias, inconsistency, indirectness, imprecision and publication bias. Downgrading the certainty of evidence for the included study outcomes was related to issues concerning the risk of bias (e.g. reported randomisation; blinding of outcomes; incomplete outcome data) as well as imprecision (e.g. wide CI, total number of participants lower than 400). We downgraded up to a maximum of three levels. We presented these ratings in the summary of findings table and provided our reasons for downgrading the certainty of the evidence in the explanations.

RESULTS

Description of studies

Results of the search

The electronic searches for this update identified a total of 1356 records (see Figure 1). These were imported in EndNote where 355 duplicates were identified, leaving 1001 records from electronic searches. Seven additional records were identified through other sources, so that 1008 records needed to be screened. We used Cochrane's Screen4Me workflow to help screen the 1001 records from the electronic searches. First, we identified 28 database records of reviews or systematic reviews which we separated from the rest of the records. The remaining 973 records from electronic searches were classified using Cochrane's Screen4Me workflow to help identify potential reports of randomised trials. The results of the Screen4Me assessment process can be seen in Figure 2 (July 2020 search) and Figure 3 (August 2021 search). We excluded 321 records as they were ineligible regarding study type (267 records when applying the Sceen4Me workflow on the results of the July 2020 search, and 45 records following the August 2021 search). Based on title and abstract assessment, we then screened the remaining 652 records left in after Screen4Me and the seven records identified through other sources, and excluded 624 (July 2020: 509; August 2021: 115). We examined the remaining 35 records in full text, and excluded nine (Bringas 2015; Cowan 2016; Dezfoolian 2013; Gooding 2011; Iseri 2014; Kim 2000; Mendelson 2016; Sanglakh Goochan Atigh 2017; Yoo 2018; see Characteristics of excluded studies). Six of these were excluded because they were not RCTs or CCTs; one because participants were not diagnosed with ASD; one because the intervention was not music therapy; and one because no relevant comparison condition was included. Additionally, four relevant ongoing studies were identified, and another ongoing study is still awaiting classification. Thus, we included 16 new studies from 20 reports, along with 10 studies from the previous update (including a new report of a previously included study, Thompson 2014), which brought the total number of included studies to 26 (see Characteristics of included studies). 25 of these studies were included in the meta-analysis. Figure 4 (a) shows the accumulation of studies over time.

Figure 1. Study flow diagram

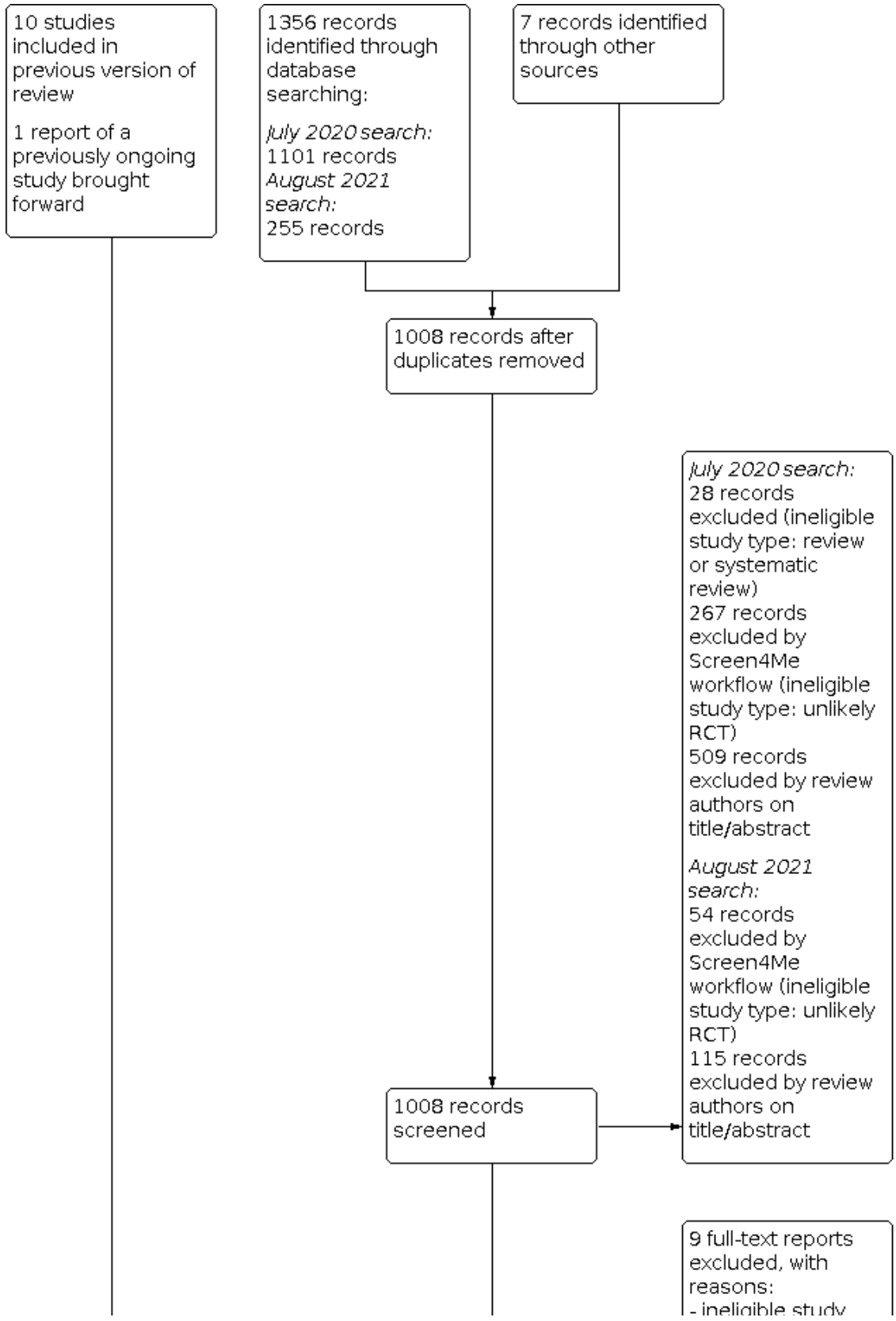


Figure 1. (Continued)

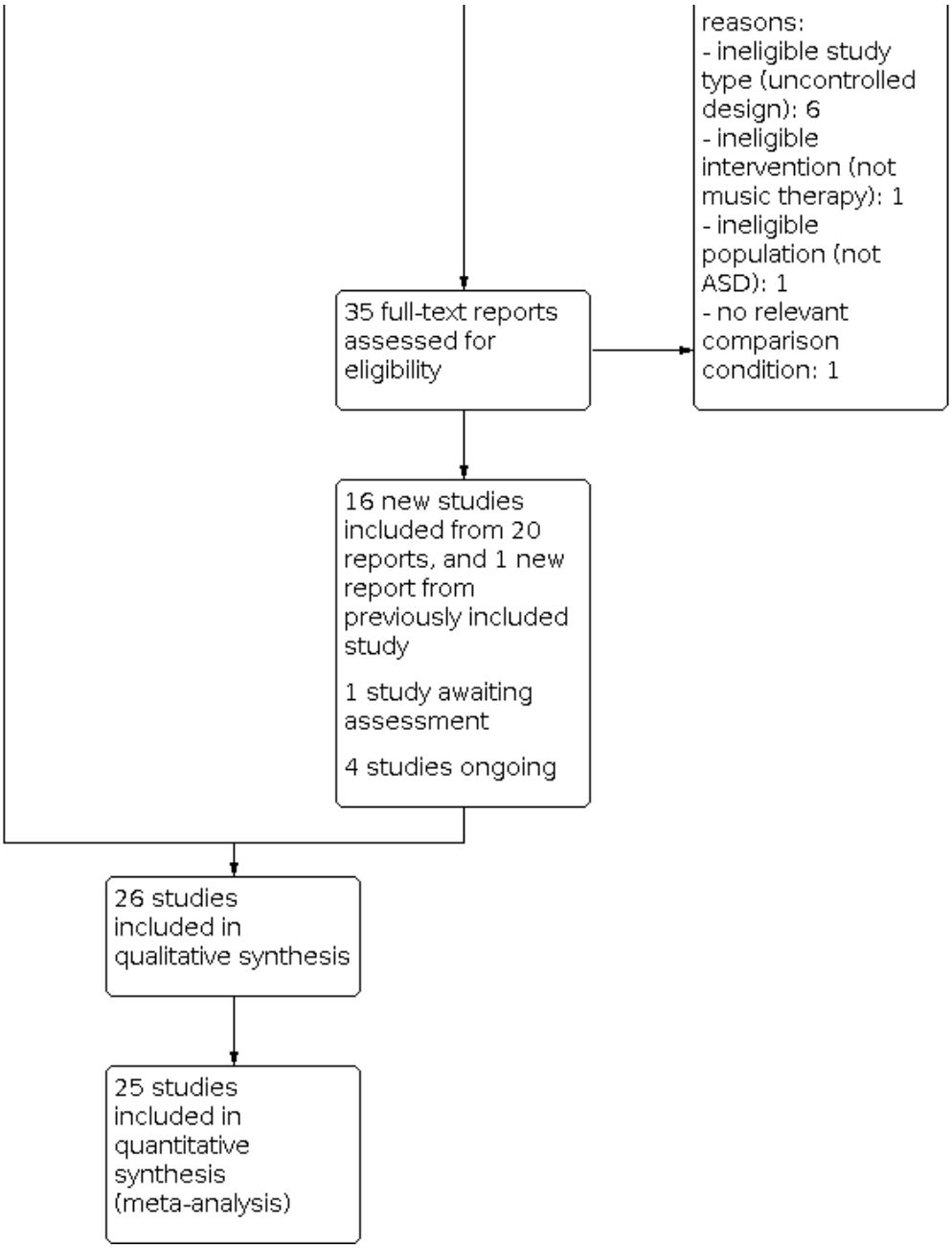


Figure 2. Sreen4Me summary diagram - July 2020 search

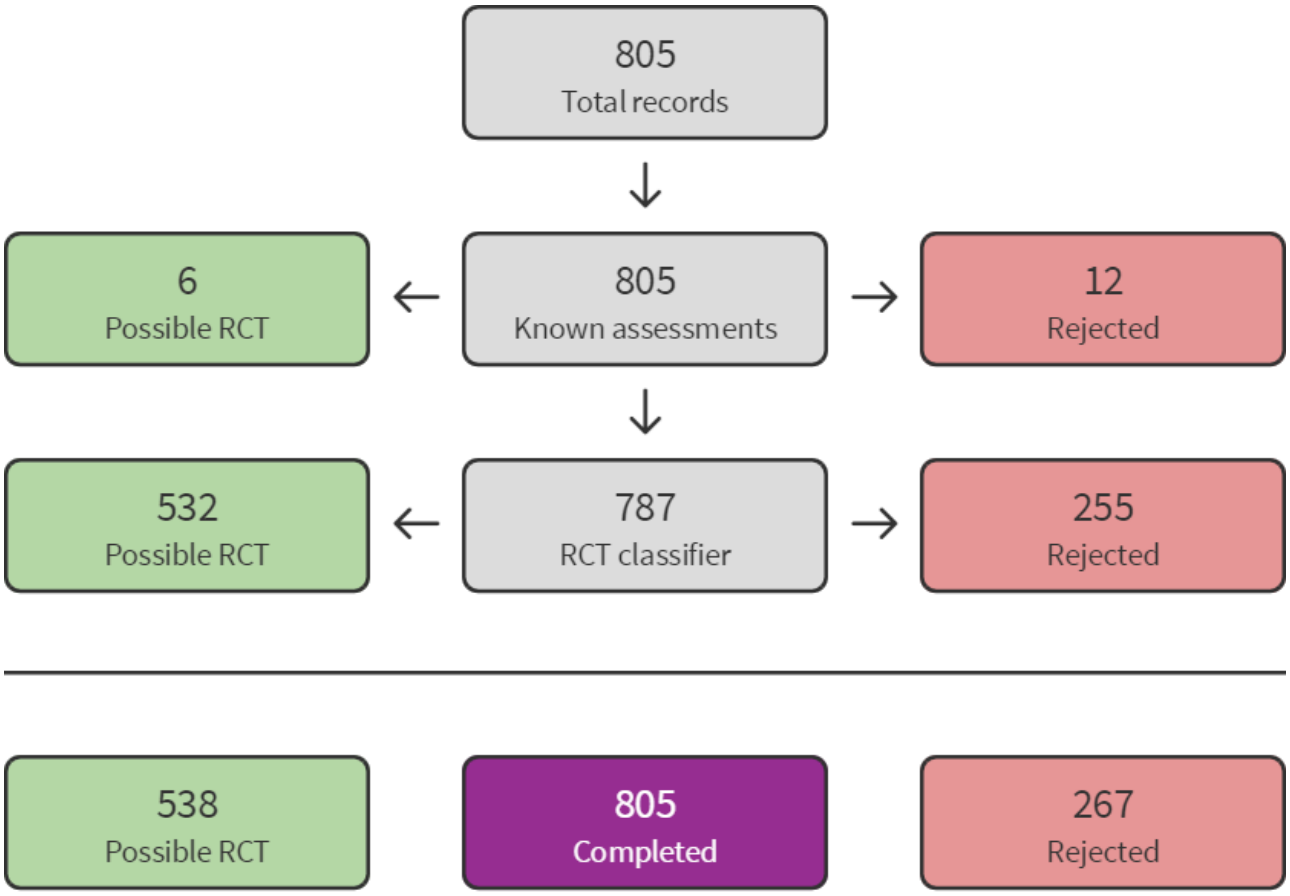


Figure 3. Screen4Me summary diagram - August 2021 search

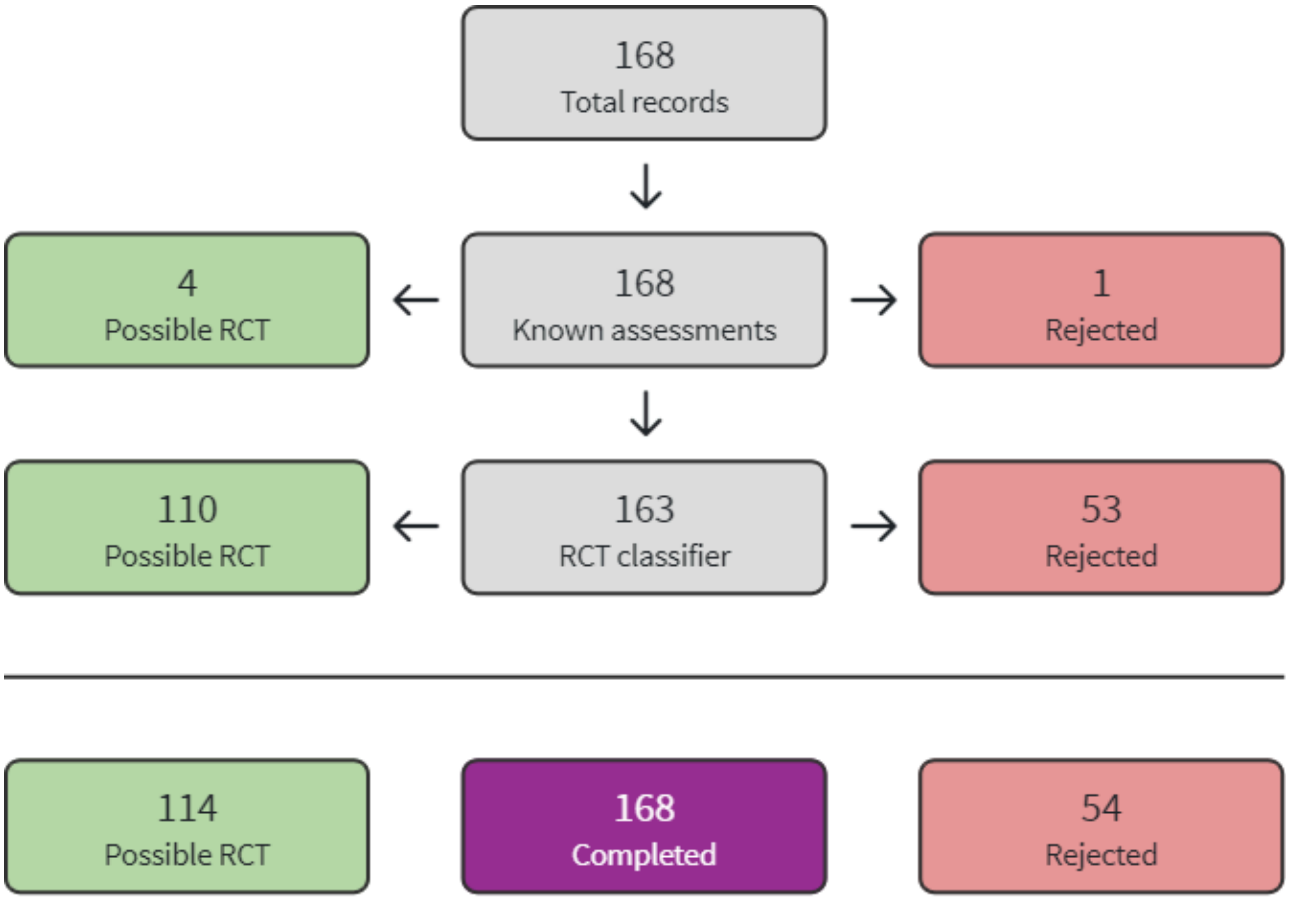
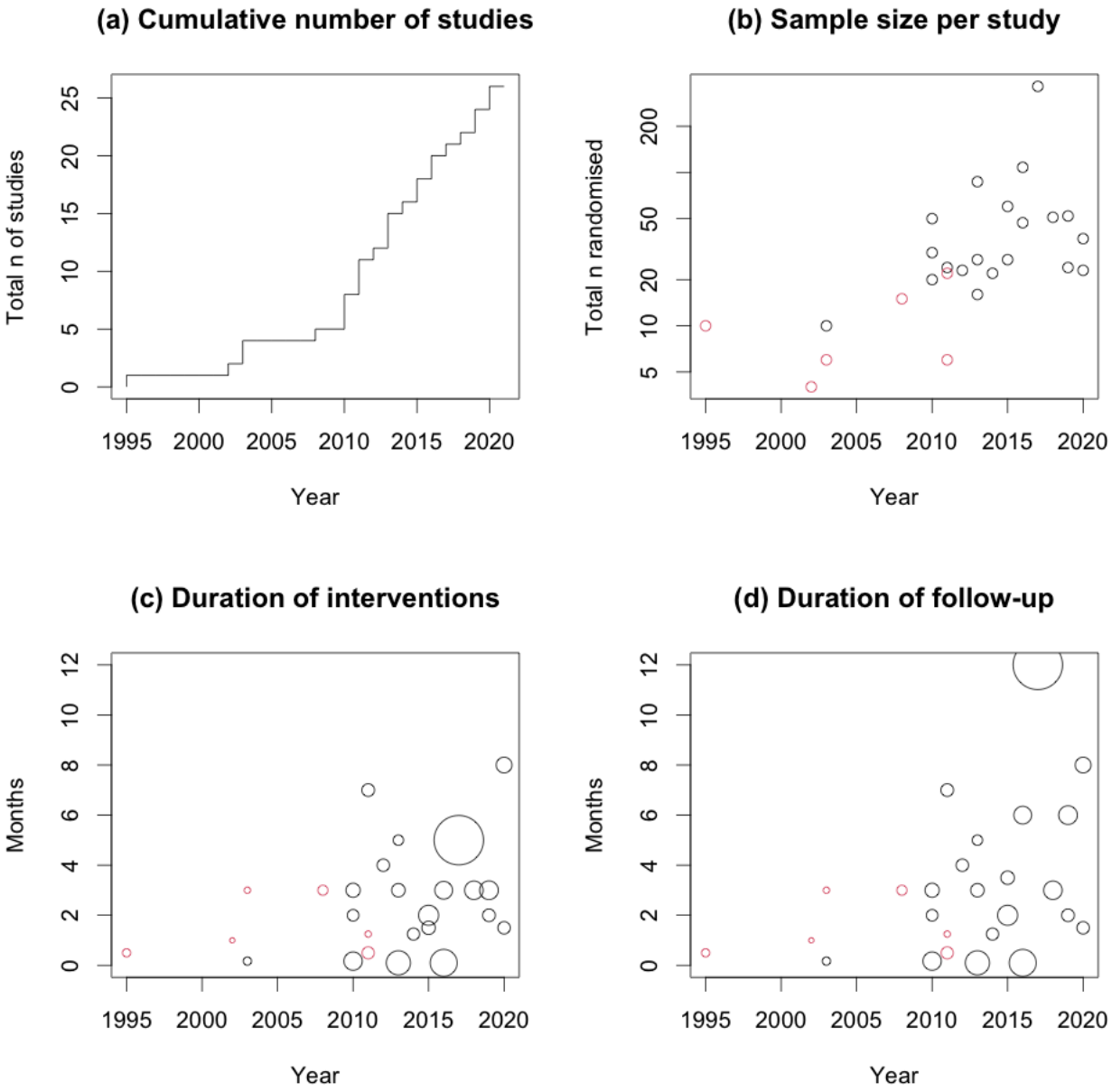


Figure 4. Accumulation of evidence from 1995 to 2020. Key: black circles = parallel design; red circles = cross-over design. Bubble sizes in panels (c) and (d) reflect number of participants randomised.



Included studies

Twenty-six studies met the criteria for the review (see Characteristics of included studies). Of these, three studies were included in the first version of this review in 2006, seven studies were added for the update of 2014, and 16 new studies (from 20 reports) were added for the present update (see Table 1 for details on this and on further summarised characteristics of included studies).

Most studies (n = 12) were conducted in North America, of which 11 were in the USA and one, Sharda 2018, in Canada. Seven studies were conducted in Asia, specifically three in China (Chen 2010; Chen 2013; Huang 2015), two in Korea (Kim 2008; Moon 2010), one in India (Bharathi 2019), and one in Iran (Ghasemtabar 2015). Four studies were conducted in Europe, i.e. France (Rabeyron 2020), Spain (Mateos-Moreno 2013), Turkey (Yurteri 2019), and the UK (Porter 2017). One study was conducted in Brazil (Gattino 2011) and one in Australia (Thompson 2014). Finally, one study (Bieleninik 2017) was a multinational trial that recruited participants in nine different countries across the world (Australia, Austria, Brazil, Korea, Israel, Italy, Norway, UK, USA).

IPD were available for 14 studies, either published or from correspondence with authors.

Length of trials

The mean duration of follow-up was 3.0 months (SD = 2.87; median 2.5; range 3 days to 12 months). The mean duration of the intervention was 2.5 months (SD = 2.05; median 2 months; range 3 days to 8 months). Figure 4 (c) and (d) shows the duration of interventions and follow-up, respectively. It can be seen that most studies lasted up to about six months (cross-over trials up to three months).

Participants

Age

Most studies (n = 21) included only children aged between two and 12 years. One study, Porter 2017, included children and adolescents, with ages ranging between eight and 16 years. Another study, Sa 2020, recruited students aged 10 to 14, but this study's data were not used in the meta-analyses. Two studies recruited both children and adults who were between nine and 21 years old (Schwartzberg 2013; Schwartzberg 2016). Finally, the study by Mateos-Moreno 2013 included only adults, with a mean age of 25 years. The majority of the participants were males (range 50 to 100%).

Diagnosis

All participants had received a diagnosis of ASD according to current or past classification systems (ICD and DSM), whether identified by a psychological assessment or a psychiatric diagnosis. The study by Porter 2017 included participants with different diagnoses (i.e. anxiety, depression, or ASD); however, only participants with an ASD diagnosis were included in the meta-analyses.

Standardised tools for diagnosis were used in eight studies (Bieleninik 2017; Ghasemtabar 2015; Gattino 2011; Kim 2008; Mateos-Moreno 2013; Rabeyron 2020; Sa 2020; Sharda 2018). Specifically, the Autism Diagnostic Observation Schedule (ADOS; Lord 1999) was used in three studies (Bieleninik 2017; Kim 2008; Sharda 2018) for diagnostic confirmation. Of these, two studies (Bieleninik 2017; Sharda 2018) used the Autism Diagnostic Interview-Revised (ADI-R; Lord 1994) in addition to the ADOS. The Childhood Autism Rating Scale (CARS; Schopler 1980) was adopted in five studies as a diagnostic tool (Gattino 2011; Ghasemtabar 2015; Kim 2008; Rabeyron 2020; Sharda 2018). The High-Functioning Version of the Childhood Autism Rating Scale (CARS2-HF; Schopler 2010) was used in Sa 2020. In the Mateos-Moreno 2013 study, the diagnosis of ASD was confirmed using the Structured Clinical Interview for DSM IV Axis I Disorders (SCID-I; First 2004). Three studies (Buday 1995; Lim 2010; Lim 2011) reported that the ASD diagnoses were performed by healthcare providers of participants. LaGasse 2014 included participants with 'a formal documentation of ASD'.

With a few exceptions (Brownell 2002; Mateos-Moreno 2013; Rabeyron 2020; Sharda 2018), the studies included both non-verbal and verbal children with varied cognitive and adaptive abilities, ranging from mild to severe autism. Brownell 2002 recruited four verbal children with 'at least prereading skills'. The Mateos-Moreno 2013 study included only young adults with severe autism. Rabeyron 2020 reported that all participants had an IQ below 70. Conversely, Sharda 2018 included only participants without intellectual disability (ID), although it was reported that 13 participants in the music therapy group had associated language impairments.

Intelligence quotient (IQ) was reported only in four studies, and was evaluated using different instruments. Bieleninik 2017 used either the Kaufman Assessment Battery for Children (KABC; Kaufman 1987), other instruments, or clinical judgement, with 45% of the sample having an IQ < 70. Gattino 2011 adopted the Raven's Coloured Progressive Matrices as a cognitive measure in 22 participants (Pasquali 2002), with six having ID. Two trials used the Wechsler scales in line with participants' chronological age: Sharda 2018 used the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler 1999) or the Wechsler Intelligence Scale for Children (WISC; Wechsler 1949), while Rabeyron 2020 used the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler 1967). Finally, Buday 1995 reported participants to be ranging from mildly to severely mentally retarded (according to DSM III-R), but did not systematically evaluate the IQ of participants.

Autism severity

Severity levels were reported in 14 studies, ranging from mild to severe autism, and were mostly evaluated using the CARS (Bharathi 2019; Buday 1995; Chen 2010; Gattino 2011; Ghasemtabar 2015; Kim 2008; LaGasse 2014; Lim 2010; Mateos-Moreno 2013; Rabeyron 2020; Sa 2020; Sharda 2018). Levels of functioning and adaptive abilities at baseline were systematically assessed only in four studies: Sharda 2018 and Thompson 2014 used different versions of the Vineland Scales (Sparrow 1998; Sparrow 1984); Chen 2013 and Kim 2008 used the Psychoeducational Profile (PEP; Schopler 1979).

Setting

The participants received therapy either at home (Thompson 2014), at school (Brownell 2002; Buday 1995; Sa 2020), in hospital (Chen 2010; Chen 2013; Gattino 2011; Huang 2015; Moon 2010), at outpatient therapy centres (Bieleninik 2017; Ghasemtabar 2015; Kim 2008; Mateos-Moreno 2013; Porter 2017; Rabeyron 2020), or a combination thereof (Farmer 2003; Lim 2010). Two studies were conducted during summer camps (Schwartzberg 2013; Schwartzberg 2016). For the remaining seven studies, the therapy setting was not reported.

Study size and design

The present systematic review involved a total of 1165 participants, with sample size ranging from 4 (Brownell 2002) to 364 (Bieleninik 2017). The median sample size was 24 participants (M = 45, SD = 70).

Twenty trials adopted a parallel design, of which two were cluster-randomised (Schwartzberg 2013; Schwartzberg 2016). Six studies had a cross-over design (Arezina 2011; Brownell 2002; Buday 1995; Kim 2008; Lim 2011; Thomas 2003). The high proportion of parallel designs is in contrast to the previous update, where the majority of included trials used cross-over designs. The cross-over trials included in this update were designed to compensate for small sample sizes: the cross-over trials ranged from four to 22 participants, whereas the parallel trials ranged from 10 to 364 participants. From Figure 4 (b) it can be seen that the sample size of studies tended to increase over time, especially in parallel trials.

Interventions

Music therapy

The majority of studies included in this review examined music therapy in an individual (i.e. one-to-one) setting (n = 13). In eight trials, music therapy was delivered in a group setting (Bharathi 2019; Ghasemtabar 2015; LaGasse 2014; Mateos-Moreno 2013; Rabeyron 2020; Sa 2020; Schwartzberg 2013; Schwartzberg 2016). One study reported that music therapy was delivered either individually or in small groups of up to three people (Yurteri 2019). Thompson 2014 applied a family-based setting where parents or other family members were also involved in therapy sessions. In four studies, it was unclear whether music therapy sessions were conducted in an individual or group setting (Chen 2010; Chen 2013; Huang 2015; Moon 2010).

The frequency of music therapy sessions ranged from daily to weekly. In seven studies music therapy was provided daily, all with a very short duration of one or two weeks. Of the studies that provided music therapy over a longer time period, it was provided weekly in nine studies, twice weekly in six studies, and in the remaining studies three (Bharathi 2019), four (Chen 2010), or six times (Huang 2015) per week. One study (Bieleninik 2017) randomised to either one or three sessions per week. The duration of sessions ranged from 10 (Arezina 2011; Lim 2010) to 60 minutes (Ghasemtabar 2015; Mateos-Moreno 2013) with a median of 30 minutes.

In two studies, we combined the data from the two music therapy groups (i.e. low-intensity and high-intensity music therapy in Bieleninik 2017; social stories music therapy and music therapy without lyrics in Chen 2013).

Content of the intervention

Twelve studies utilised a highly structured approach to music therapy using receptive techniques (i.e. listening to live or, in the case of Lim 2010 and Rabeyron 2020, pre-recorded music presented by the therapist) or a mix of receptive elements and active music-making. Songs sung by the music therapist were composed or chosen individually for the participants and were usually used with specific aims. For example, songs were based on a social story addressing a central problem behaviour of the particular individual in treatment (Brownell 2002) or autistic individuals in general (Schwartzberg 2013; Schwartzberg 2016); they contained signs and words to be learned (Buday 1995; Lim 2010; Lim 2011); or they were used to build a relationship and to provide a safe and understandable structure for the participants in the study (Chen 2010, Chen 2013, Farmer 2003). Active music-making by the participants, which is often typical for music therapy in clinical practice (Wigram 2006), was reported in five of those studies (Chen 2010; Chen 2013; Farmer 2003; Moon 2010; Sa 2020). Participants were invited to play guitar, pitched or unpitched percussion instruments, and sing songs. Playing instruments was partly used to reinforce adjusted behaviour. Moon 2010 used a music drama based on a theory of mind approach, including narration, singing, and musical instrument playing.

In the other fourteen studies, particular emphasis was put on the interactive and relational aspects of music therapy. Music therapy techniques included improvisation, songs, and structured musical games. Interventions followed a non-directive approach and focused on engaging the individual in musical interaction, offering opportunities for the individual to make choices and to initiate contact. Generally, the therapist's interventions were depicted as drawing on the individual person's skills, interests, preferences, and motivations as well as on their immediate expression and behaviour. By attuning to the individual musically and emotionally, the therapists create moments of synchronisation that help the individual to experience and recognise core elements of reciprocal communication (Kim 2008; Schumacher 1999a; Schumacher 1999b; Stephens 2008; Thompson 2014; Wigram 2009). Mateos-Moreno 2013 combined music therapy with dance/movement therapy activities, such as massages with small balls, simulation situations, imitation, role-playing, and dancing.

Several of the studies employed specifically developed treatment guidelines in the form of a treatment contingency plan (Thompson 2014), or a treatment manual (Bieleninik 2017; Ghasemtabar 2015; Kim 2008; LaGasse 2014; Porter 2017; Sa 2020; Sharda 2018). In these protocols, principles and procedures of therapy are specified whilst allowing the therapist to adapt interventions flexibly according to the child's needs and the specific requirements of the situation.

Comparators

'Placebo' therapy

A total of 15 studies used a 'placebo' therapy to control for the non-specific elements of the therapy. Thirteen of these used a 'placebo' activity to control for the non-specific effects of therapeutic attention. Since, in all of these studies, music was considered as the specific ingredient of music therapy, the placebo conditions were constructed to closely match the music therapy condition, only that music was not used. Specifically, a social story was read instead of sung to the participants (Brownell 2002; Moon 2010; Schwartzberg 2013; Schwartzberg 2016); rhythmic or normal speech was used instead of singing (Buday 1995; Lim 2010; Lim 2011); play activities were offered without using songs or musical instruments (Farmer 2003); the therapist engaged in interaction with the child by responding to the child's behaviour non-musically and using non-music toys (Arezina 2011; Kim 2008; Sharda 2018; Thomas 2003); or the participants were involved in a social skills group (LaGasse 2014). Two studies (Bharathi 2019; Rabeyron 2020) used another type of 'placebo' therapy consisting of passive music listening. In both studies participants passively listened to songs played using a CD player, without any interaction with the therapists. Thus, not the music, but the therapist's attention was seen as the specific ingredient in these studies.

Standard care

Eleven studies compared music therapy with standard care. In Bieleninik 2017, the control group received enhanced standard care, which consisted of the routine care available at the site, plus three 60-minute sessions of parent counselling over the five months of intervention (at 0, two, and 5 months). Three studies (Chen 2010; Chen 2013; Huang 2015) compared music therapy with a comprehensive/integrated treatment including several activities, such as auditory integration training, sensory integration, special education, language therapy, speech therapy, and play therapy. Gattino 2011 reported that participants received routine clinical services, including medical examinations and consultations. Sa 2020, where a waiting-list control design was applied, and Ghasemtabar 2015 described no intervention. In Mateos-Moreno 2013, participants in the control group

were attending their regular therapies as well as receiving pharmacological treatment. Analogously, in the Porter 2017 study, participants were following psychiatric counselling and/or medication. In the Thompson 2014 study, participants received varying forms of services and support from early childhood intervention centres. Finally, in Yurteri 2019, participants in the control condition received no treatment except monthly routine child psychiatric follow-up and special education.

Multiple-armed trials

Some studies included other conditions whose data were not included in this review. Brownell 2002 reported observations during a baseline period and a washout period with no intervention. Arezina 2011 also observed behaviour in an 'independent play' group, which we considered was neither 'placebo' therapy nor 'standard care'. Therefore, data from this group were not included in this review. Lim 2010 and Lim 2011 compared music training with both a speech training (included) and a 'no training' group (excluded).

Outcome measures

Both generalised and non-generalised outcomes were used in the included studies. Non-generalised outcomes refer to changes in the child's non-generalised behaviour in the same setting where the intervention takes place, as opposed to generalised outcomes which are observed in other settings (Warren 2011).

Primary outcomes

1) Global improvement

Global improvement was defined as a binary outcome (improved versus not improved or unknown, on a scale measuring clinical global impressions or on a global measure used as primary outcome in a study). The negative outcome was imputed for missing values, enabling a full intention-to-treat analysis. Global improvement was measured using the Clinical Global Impression scale (CGI; Busner 2007) or if this was not available, the primary outcome chosen by study authors.

Rabeyron 2020 used the Clinical Global Impression-Severity scale (CGI-S; Busner 2007). The CGI-S is a 7-point clinician-rated scale used to rate the severity of a disorder, with higher scores indicating greater severity. The scores range between 1 ('normal') and 7 ('among the most extremely ill patients').
Seven studies (Bieleninik 2017; Bharathi 2019; Kim 2008; LaGasse 2014; Porter 2017; Schwartzberg 2013; Thompson 2014) had a clearly defined primary outcome (other than CGI) and provided IPD from which to calculate global improvement.

2) Social interaction

Social interaction was examined in 14 studies. The following scales were used:

i. The TRIAD Social Skills Assessment (TSSA; Stone 2010) is a 'criterion-based tool' which provides specific assessment considering parent, teacher, observation, and direct interaction with the children aged six to 12 years. It consists of three components: Problem Behavior Rating Scale, Social Skills Survey, and Social Skills Rating Form. In Bharathi 2019, one of the three components of the TSSA (i.e. the Social Skills Rating Form) was used. Each item is rated on a 4-point Likert scale, with higher scores indicating more favourable behaviours.

ii. The 'social communication' domain of the Childhood Autism Rating Scale (CARS; Magyar 2007; Brazilian version: Pereira 2008; Rapin 2008) was used in three studies (Chen 2013; Gattino 2011; Rabeyron 2020). The CARS (Schopler 1980) is a 15-item observation-based behavioural rating scale administered by health professionals for the diagnosis of children with autism and pervasive developmental disorders. Total scores can range between 15 and 60, with higher scores indicating higher severity. The 'social communication' domain has been derived from the factor analysis of the CARS (see Magyar 2007; McConachie 2015) and is composed of five items of the original tool, all related to social communication skills (i.e. imitation, verbal and nonverbal communication, consistency of intellectual responses and general impressions). Similarly to the full scale, this domain was administered by investigators blind to group allocation (unclear for Chen 2013). As in Chen 2013 SD were missing, we imputed SD = 3 in line with other studies using the same scale.

iii. The 'social affect' (SA) subscale of the Autism Diagnostic Observation Schedule (ADOS; Lord 1999) was used in Bieleninik 2017. The ADOS is a semi-structured, interactive observation by trained health professionals. It has been designed to assess aspects of communication, social reciprocal interaction, play, and stereotyped behaviours and restricted interests. It consists of four modules, appropriate for individuals with different developmental and language levels. ADOS-SA is composed of two subdomains, i.e. 'language and Communication' and 'reciprocal Social interaction'. The ADOS-SA score can range from 0 to 24 (module 1 and 2) or 0 to 27 (module 3), with higher scores indicating greater symptom severity. In the study by Bieleninik 2017, it was rated by independent, blinded health professionals.

iv. The total score of the Social Responsiveness Scale (SRS; Constantino 2005) was used in four studies (Bieleninik 2017; LaGasse 2014; Sharda 2018; Thompson 2014). The SRS is a 65-item scale measuring the severity of autism symptoms as they occur in natural social settings. The total score can range from 0 to 195. Higher scores are indicative of greater symptom severity. The SRS is rated by parents or teachers and it is appropriate for use with children from four to 18 years of age. In Thompson 2014, the Preschool Version of the SRS was used. Sharda 2018 used the SRS-2, a revised and more recent version of the SRS (Constantino 2012). For Sharda 2018, where the SD was not reported, we imputed SD = 30 based on other studies using the same scale.

v. The Social Skills Rating System scale (SSRS; Gresham 1990), elementary form, was completed by participants' parents in Ghasemtabar 2015. The total score can range between 0 and 80. Higher scores indicate higher social skills and thus favourable outcome.

vi. The 'social approach behaviours' subscale of the Pervasive Developmental Disorder Behavior Inventory, Korean version (PDDBI; Cohen 1999) was used in Kim 2008. The scale was filled out by professionals (i.e. a teacher or a therapist of the child) who were blind to experimental condition. Higher scores are indicative of better social skills.

vii. The total score of the Social Skills Improvement System (SSIS) rating scales (Gresham 2008) was used in one study (Porter 2017). The scale was rated by parents and self-rated by youth. The SSIS is a scale with 75 (self) to 79 (parent) items across 3 subdomains (i.e. social skills, competing problem behaviours,

academic competence). The total score can range from 0 to 225 (self) or 237 (parent), with higher scores representing favourable outcomes.

viii. The Autism Social Skills Profile (ASSP; Bellini 2007) was used in Schwartzberg 2013. The ASSP is a 49-item tool divided into three sub-categories: social reciprocity (SR), composed of 23 items; social participation (SP), composed of 12 items; and detrimental social behaviours (DSB), composed of 10 items. Each item is rated on a 4-point Likert scale. Even though Schwartzberg 2013 calculated the ASSP score for each sub-category, the ASSP total score was used for the outcome 'social interaction'. The scale was completed by participants' legal guardians.

ix. The Vineland Social-Emotional Early Childhood Scales (SEEC; Sparrow 1998) were used in one study (Thompson 2014). The Vineland SEEC is a 88-item measure used to assess the social and emotional functioning of children from birth through 5.11 years. In Thompson 2014, It was administered through a semi-structured interview with the child's parent participating in the study. Only two out of three subscales (i.e. interpersonal relationships; play and leisure time) were used in the study.

Some studies used more than one score to measure social interaction at the same time point. Bieleninik 2017 used both ADOS-SA and SRS; only ADOS assessors were blinded, however both perspectives of parents (SRS) and professionals (ADOS-SA) were important so we merged both. Porter 2017 used parent and self-reports of the same scale; both were considered equally valid; we merged both to represent both perspectives. Thompson 2014 used both the SRS and Vineland SEEC; again, we merged both because both were equally valid.

For the meta-analysis, given that some scales in this domain were 'negative' (ADOS, SRS, CARS) and others 'positive' (ASSP, SSIS, SSRS), we reversed the 'negative' scales in the analysis so that the positive sign of the analysis matches the positive meaning conveyed by 'social interaction' (i.e. positive effects represent a favourable outcome).

Three studies assessed social interaction skills using non-validated outcome measures, through the observation of participants' behaviour within therapy sessions:

i. In Arezina 2011, the researcher coded videotaped sessions for 'requesting (initiating joint attention)' behaviours such as pointing, giving an object to the therapist, or touching the therapist while making eye contact; an independent observer additionally coded a third of the session material. In Thomas 2003, 'requesting behavior' was defined in a similar way. Video tapes were coded by a music therapy intern and rated for two outcomes, task behaviour and requesting.

ii. One study (Kim 2008) also investigated observed behaviours related to social interaction in the intervention setting. These measures included frequency and duration of the child's turn-taking, frequency of imitation behaviours, frequency and duration of both 'emotional synchronicity' and 'musical synchronicity', and behaviours associated with the frequency and duration of joy (i.e. smiling and laughing) on the part of the child. The coding procedure was conducted by the lead investigator by microanalytically (second by second) observing DVD recordings, with subsequent coding supplemented by a trained research assistant who was blind to session order.

3) Non-verbal communication

Non-verbal (i.e. gaze-related and gestural) communication was examined in 11 studies. Six studies used validated outcome measures, as follows:

i. The 'non-verbal communication' domain of the CARS was used in three studies (Chen 2013; Gattino 2011; Rabeyron 2020).

ii. The Early Social Communication Scales (ESCS; Mundy 2003) is a videotaped structured play-based assessment measuring non-verbal social communication skills in children aged between six and 30 months. In Kim 2008, the shortened version of the ESCS was used. The ESCS provides frequencies of scores for 'initiation of joint attention' and 'responding to joint attention'. The scoring was administered by the researcher and by two trained research assistants who were blind to group assignment.

iii. The Children's Communication Checklist (CCC-2; Bishop 1998), used in Sharda 2018, is a parent/caregiver-administered 70-item rating scale to measure children's social communication skills across 10 domains. This tool is focused on the assessment of non-verbal communication, pragmatics, as well as aspects of language structure and discourse. Sharda 2018 used the standard general communication composite standard score as a measure of the child's general pragmatics and communication ability. Higher scores indicate better social communication skills.

iv. The MacArthur-Bates Communicative Development Inventories – Words and Gestures (MBCDI-W&G; Fenson 2007) are a set of parent-rated measures designed to evaluate the verbal and non-verbal communicative skills of young children. The section 'action and gestures' of the MBCDI-W&G was used as a measure of non-verbal communication in Thompson 2014. Higher scores are indicative of higher levels of non-verbal communication.

Four studies assessed non-verbal communication skills using non-validated outcome measures, through the observation of participants' behaviour within therapy sessions (Buday 1995; Farmer 2003; Kim 2008; LaGasse 2014). Measures of non-verbal communication skills in these studies are reported below:

i. In Buday 1995, the outcome consisted simply of the number of signs correctly imitated within a session.

ii. In Farmer 2003, a completed gesture was given a score of two, and an attempt a score of one, and the outcome consisted of the sum of these scores for all attempted and completed gestures within a session.

iii. In Kim 2008, frequency and duration of eye contact (i.e. the child looking at the therapist) was coded by microanalytic analysis of the session material.

iv. In LaGasse 2014, video recordings of children in both groups were analysed for instances of group communication and social interaction attempts. Two trained music therapy research assistants completed the coding of predefined behaviours (i.e., eye gaze, joint attention, initiation of communication, response to communication, withdrawal behaviours). Five-minute

clips were randomly selected from each session for each child. The session order was concealed from the coders.

4) Verbal communication

Communicative skills in verbal communication were addressed in 11 studies. The authors used the following outcome measures:

i. The 'verbal communication' domain of the CARS was used in three studies (Chen 2013; Gattino 2011; Rabeyron 2020).

ii. Thompson 2014 used the subscales 'phrases understood', 'words understood' and 'words produced' of the MBCDI-W&G (Fenson 2007).

iii. Comprehension Checks (CCs) were used by Schwartzberg 2013 and Schwartzberg 2016. They consisted of a series of five close-ended questions (yes or no) to evaluate participants' comprehension of social stories.

iv. The Peabody Picture Vocabulary Test (PPVT-4; Dunn 1981), a short, standardised measure of one-word receptive vocabulary, was used in one study (Sharda 2018). The test requires the participant to choose one of four colour pictures on a page. Higher scores indicate better receptive vocabulary.

v. For Buday 1995, Farmer 2003, Lim 2010, and Lim 2011, independent observers rated in-session behaviour by counting the frequency of appropriate verbal responses in a manner similar to the previous outcome. The outcome measures used in these four studies were unpublished.

5) Quality of life

Quality of life (QoL) was measured in three studies, using three different scales:

i. Bieleninik 2017 evaluated the QoL of both the child and the family as a whole using a Visual Analogue Scale (VAS) ranging from 0 to 100, where 0 corresponded to the worst and 100 to the best possible QoL.

ii. Sharda 2018 used the Beach Center Family Quality of Life Scale (FQoL; Park 2003) to assess satisfaction with different aspects of family quality of life. FQoL is a 25-item questionnaire containing five subscales: family interaction, parenting, emotional well-being, physical/material well-being, and disability-related support. Higher scores correspond to better satisfaction in QoL.

iii. Yurteri 2019 evaluated participants' QoL using the Pediatric Quality of Life Inventory (PedsQL; Varni 1999). It consists of a 23-item scale designed to measure the core dimensions of health as delineated by the World Health Organization, as well as role (school) functioning. In Yurteri 2019, the scale was completed by parents according to participants' age. The PedsQL is a multidimensional scale composed of four dimensions (i.e. physical functioning, emotional functioning, social functioning, school functioning). Moreover, three summary scores can be calculated (i.e. total scale score, physical health summary score, psychosocial health summary score). The Yurteri 2019 paper reported both the total scale and the psychosocial health summary scores. Higher scores correspond to better quality of life.

6) Total autism symptom severity

Total autism symptom severity was measured in nine studies. Outcome measures included the following:

i. The CARS (Schopler 1980) was used in three studies (Bharathi 2019; Chen 2013; Rabeyron 2020), although in Chen 2013 CARS scores were not reported or made available to us.

ii. The total score of the ADOS (Lord 1999) was used by Bieleninik 2017. The ADOS total score is calculated summing up the raw scores of the ADOS-SA and 'restricted and repetitive behaviour' (ADOS-RRB) scores.

iii. the Autism Treatment Evaluation Checklist (ATEC; Rimland 1999) was used by LaGasse 2014. The ATEC is a 77-item checklist that includes four areas (speech and communication, sociability, sensory/cognitive awareness, health/physical behaviour). It is completed by parents, teachers and/or primary caretakers of autistic children. Total ATEC scores range from 0 to 180. Lower scores on ATEC demonstrate higher functioning.

iii. The total score of the Autism Behavior Checklist (AuBC; Krug 1980) was adopted as a measure of total symptom severity in four studies (Chen 2010; Chen 2013; Huang 2015; Yurteri 2019). It consists of 57 items with higher scores indicating higher severity. For Huang 2015, where no SD was reported, we imputed SD = 12 from other studies that used the AuBC (Chen 2010; Chen 2013; Yurteri 2019).

v. The Revised Clinical Scale for the Evaluation of Autistic Behavior (ECA-R; Barthélémy 2003) was used by one study (Mateos-Moreno 2013). It is composed of 29 items with lower scores corresponding to favourable outcomes.

7) Adverse events

Two studies collected adverse event data. In Bieleninik 2017, hospitalisation or other institutional stay (including pre-planned stays) were included as adverse events; these and any other serious or non-serious adverse events were reported by parents. Porter 2017 collected serious adverse events and non-serious adverse events related to study procedures. None of the other studies reported information on adverse events.

Secondary outcomes

8) Adaptive behaviour

Nine studies evaluated adaptive behaviours. Validated scales were used in five studies, along with other non-validated measures:

i. The Psychoeducational Profile (PEP; Schopler 1979; Muris 1997) was used by Chen 2010. The PEP consists of a series of toys, objects, and games which are offered to the child. It provides information on developmental items and pathology items. Higher scores are favourable. In the Chen 2010 study, a total score as well as the scores of three domains (i.e. relationship and emotions, interest in games and objects, sensory response) were provided. The total scores were used as measure of adaptive behaviour.

ii. The Child Behavior Checklist (CBC; Achenbach 2001), used by Porter 2017, is a parent-rated tool consisting of 113 questions, scored on a three-point Likert scale (0: absent, 1: occurs sometimes, 2: occurs often). Thus, lower scores are favourable. CBC scores were

not reported in the publication, but available in IPD from the study authors.

iii. The Aberrant Behavior Checklist (ABC; Aman 1985) is a 58-item caregiver-report checklist designed to assess maladaptive behaviours in people with developmental disabilities. Higher scores correspond to greater maladaptive behaviours. The ABC Total Score was used in Rabeyron 2020.

iv. The maladaptive behaviours subdomain of the Vineland Adaptive Behavior Scales (VABS; Sparrow 1984) was used by Sharda 2018 to identify the presence of behavioural problems, such as challenging internalising and externalising behaviours. The scale is administered as a semi-structured interview to a parent or caregiver. Lower scores are favourable.

v. Three studies investigated adaptive behaviours within the interventions setting (Arezina 2011; Kim 2008; Thomas 2003). In Arezina 2011 and Thomas 2003, videotaped sessions were coded for 'interaction (engaging in joint attention)' and 'on-task behavior', respectively; this included activities such as following a direction, physically manipulating a toy in a functional manner, and imitating a movement or vocal sound. In Kim 2008, sessions were scored by frequencies of 'compliant response', 'non-compliant response', and 'no response'.

vi. Restricted and repetitive behaviours were measured in Bieleninik 2017 using the ADOS-RRB domain (Lord 1999). Higher scores indicate more severe repetitive behaviours.

vii. Brownell 2002 addressed occurrence of individually targeted repetitive behaviours outside therapy sessions. Independent observers (i.e. teachers) counted how often the targeted behaviour occurred in the classroom. The frequency count was used as the outcome measure. No published scale was used in the Brownell 2002 study.

Where necessary, we reversed scores so that a high score on adaptive behaviour indicated a favourable outcome.

9) Quality of family relationships

Family relationships were evaluated in three studies, with different tools:

i. Kim 2008 used the Mother Play Intervention Profile (MPIP), a measure specifically developed for the study to describe characteristics of interactions between mothers and autistic children during a casual play situation at their home. Scores were based on video observations conducted by the researcher, supplemented by an independent observer's coding for a third of the sessions.

ii. In Porter 2017, the McMaster Family Assessment Device (FAD; Epstein 1983) was completed by parents. The FAD is a 60-item questionnaire that measures an individual's perceptions of his/her family. Each item is scored on a 4-point scale. The higher the score, the more problematic the family member perceives the family's overall functioning

iii. Thompson 2014 used the Parent-Child Relationship Inventory (PCRI; Gerard 1994), a self-report questionnaire for parents to assess the parent-child relationship and parents' attitudes towards parenting. The full instrument consists of 78 items, rated on a 4-level scale ranging from 'strongly agree' to 'strongly disagree'. Higher scores are indicative of positive parenting.

10) Identity formation

Identity formation includes all the processes that allow autistic people to develop a clear and unique view of themselves and of their identity. Domains related to identity formation were evaluated in two studies.

i. the Bandura self-efficacy scale (Bandura 1978) was used to measure self-efficacy in the Moon 2010 study. The scale is composed of nine items. Higher scores indicate higher self-efficacy levels.

ii. the Fenigstein self-awareness scale (Fenigstein 1979) was used to measure self-awareness in the Moon 2010 study. It is composed of 20 items, with higher scores indicating greater levels of self-awareness.

iii. the Rosenberg self-esteem scale (Rosenberg 1965) was adopted as a measure of self-esteem in both Moon 2010 and Porter 2017. It is a 10-item self-report scale that measures global self-worth by measuring both positive and negative feelings about oneself. All items are answered using a 4-point Likert scale format ranging from 'strongly agree' to 'strongly disagree'. The total score can range between 10 and 40. Higher scores indicate higher self-esteem.

11) Depression

Depression was evaluated in one study (Porter 2017), using the Center for Epidemiological Studies Depression Scale for Children (CES-DC; Faulstich 1986; Weissman 1980). This is a 20-item self-report questionnaire for young people between the ages of six and 17. It asks young people to rate how many depressive symptoms they have experienced in the last week. Higher scores represent higher levels of symptoms.

12) Cognitive ability

Cognitive ability was evaluated in one study (Sa 2020), using the Test of Everyday Attention for Children 2 (TEA-Ch2; Manly 2016). The TEA-Ch2 is a tool for young people between the ages of five and 15 that assesses three areas of attention skills (selective attention, sustained attention, and attentional control/switching attention) using eight tasks. However, data from this study were not included as the outcome measure was not applied by an independent rater, but by the researcher who also administered the intervention protocol (i.e. the therapist), thus violating this review's eligibility criteria for outcome measures.

Funding sources

The American Music Therapy Association (AMTA) provided funding support for two studies (LaGasse 2014: Arthur Flagler Fultz Research Fund; Thomas 2003: Mid-Atlantic Region of the AMTA). University funding was available for two studies (Kim 2008: Aalborg University, Denmark; Thompson 2014: University of Melbourne, Australia). The Thompson 2014 study was also supported by the Victorian Department of Education and Early Childhood Development. Further funding sources included the Science and Engineering Research Board, Government of India, New Delhi (Bharathi 2019); the Chongquing Natural Science Foundation (Chen 2010; Chen 2013); the Chongqing Medical Specialty Construction (Chen 2013); the Fund of Incentive to

Research of Porto Alegre Clinical Hospital and the Brazilian Research Council (Gattino 2011); the Big Lottery Fund (Porter 2017); Entreprendre pour Aider and the Academie Francaise (Rabeyron 2020); the Canadian Institutes of Health Research and Quebec Bioimaging Network (Sharda 2018). Bieleninik 2017 was supported by the Research Council of Norway, the University of Bergen, Norway, POLYFON Knowledge Cluster for Music Therapy, and a range of further governmental and university funding sources and foundations across participating countries (see Characteristics of included studies for details). For the remaining 14 studies, no funding sources were reported, or sources of support were reported as 'nil' (Ghasemtabar 2015).

Ongoing studies

Four relevant studies were still ongoing at the time of assessment (see Ongoing studies). Conducted in the USA, NCT03560297 used a cross-over design and applied a parent-child music class programme including parent training, peer inclusion, and musical play for 12 weeks, compared with a waiting-list programme. The estimated sample size of children aged 20 to 72 months was 68. Primary outcomes included a standardised motor imitation assessment, and parent questionnaires on non-verbal communication, parenting stress, and parenting efficacy/quality.

Conducted in South Korea, ISRCTN18340173 used a parallel design involving propensity score matching and applied weekly improvisational music therapy sessions for one year in addition to standard care, compared with standard care alone. The estimated sample size of children aged 24 to 72 months was 50. Primary outcomes were the ADOS and the CARS-2.

Conducted in Hong Kong, NCT04557488 used a parallel design and applied a 12-week social skill intervention using group music therapy, compared with a 12-week non-musical intervention (i.e. behavioural-based social skill group training). The estimated sample size of children aged six to 13 years was 80. Primary outcomes included the CARS-2, the SRS-2, and in-session social behaviour.

Conducted in Austria and Norway, NCT04936048 uses a cross-over design and applied 12 weekly sessions of one-on-one music therapy with an equal number of non-musical one-on-one play therapy sessions. The estimated sample size of children aged six to 12 years was 80. Primary outcomes included the CCC-2 and measures of brain connectivity of frontotemporal regions.

Studies awaiting classification

One potentially relevant study is awaiting assessment since the information available in the trial registration was not sufficient to assess eligibility (NCT03267095); recruitment has not started. It is planned to be a randomised, unblinded study, conducted in Egypt, comparing the effects of a music therapy intervention to parent counselling over a 12-month period. The researchers planned to recruit 60 children between three and seven years old with an IQ > 75. The outcome was focused on verbal communication, through the administration of an Arabic Language test evaluating semantics, expressive morphology, syntax, and pragmatics.

Excluded studies

Nine studies identified through the update search were excluded for the following reasons: six studies did not have an RCT or CCT design (six case series, i.e. studies comparing different treatments that all participants received in the same order); one study because the intervention was not music therapy, but movement activities with music; one study because participants were not diagnosed with ASD, but with severe neurological disorders; and one study because it did not include a relevant comparison condition (both groups were music therapy). See Characteristics of excluded studies, where in addition to the nine studies excluded with reasons in this update, we also report seven studies that were excluded in previous versions of this review. From the fifty-nine studies excluded with reasons in the two previous versions of this review, these seven were selected in a process of reassessment as the most relevant that one might expect to see in this review. Six of them were excluded because they did not have an RCT or CCT design; one study because it was not an intervention, but an assessment study.

Risk of bias in included studies

A visual representation of the included studies' risk of bias for each domain, as specified below, is shown in Figure 5. Figure 6 provides a summary of the risk of bias results for each included study.

Figure 5. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies

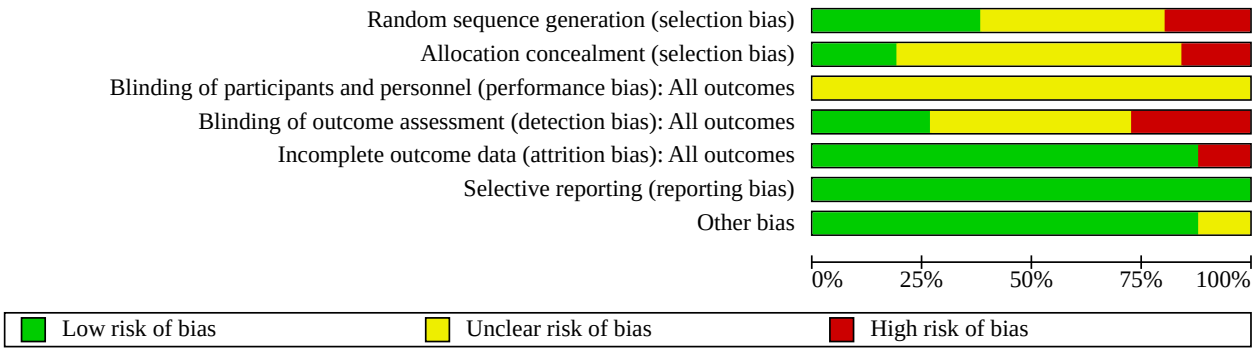


Figure 6. Risk of bias summary: review authors' judgements about each risk of bias item for each included study

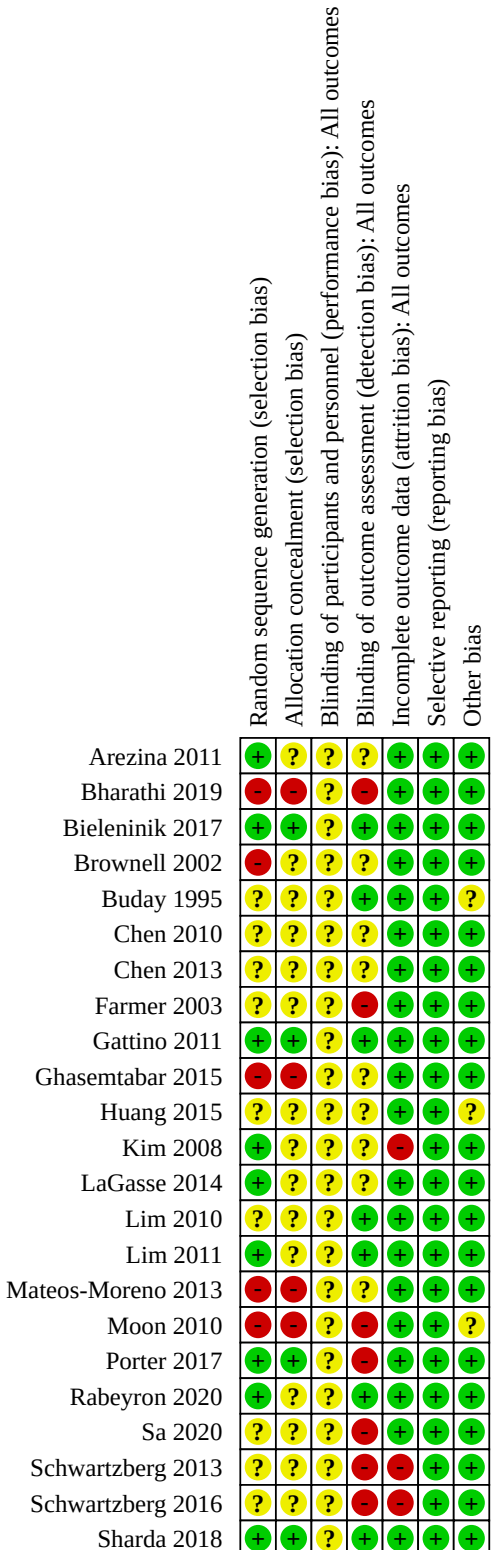
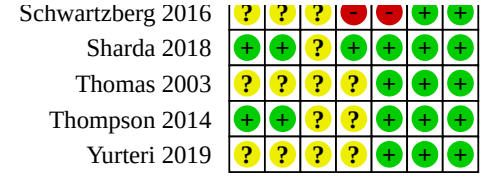


Figure 6. (Continued)



Allocation

Twenty of the included studies stated explicitly that randomisation was used to assign participants to treatment groups. Methods of randomisation included using computer-generated random sequences for determining allocation to an experimental condition (Bieleninik 2017; Gattino 2011; LaGasse 2014; Rabeyron 2020; Sharda 2018; Thompson 2014), manually generated random sequences by, for example, coin tossing (Kim 2008; Lim 2011; Sharda 2018), and a Latin Square for determining session order (Arezina 2011). We judged these studies as being at low risk of bias. In eleven of the 26 studies (Buday 1995; Chen 2010; Chen 2013; Farmer 2003; Huang 2015; Lim 2010; Thomas 2003; Sa 2020; Schwartzberg 2013; Schwartzberg 2016; Yurteri 2019), methods of randomisation were not specified and the risk of bias was judged as unclear. In the remaining five studies, no information about randomisation was provided or the described methods of randomisation did not ensure a random allocation of participants; these were rated as having high risk of bias.

The description of allocation concealment was adequately described in two studies (Bieleninik 2017; Thompson 2014) and partly clarified in three studies (Gattino 2011; Porter 2017; Sharda 2018); we judged these to be at low risk of bias. For three studies, allocation was not concealed and these were judged as being at high risk of bias (Bharathi 2019; Ghasemtabar 2015; Mateos-Moreno 2013). The remaining eighteen studies did not provide specific information on allocation concealment and the risk of bias was rated as unclear.

Blinding

Due to the nature of the intervention, it was not possible to blind those who delivered music therapy or those who received it. Consequently, neither participants nor therapists of the studies under review could be declared as blinded. However, although autistic individuals were not blinded, this was unlikely to introduce bias as they were usually not fully aware of available treatment options or study design (Cheuk 2011). The possible risk of bias introduced by therapists administering the intervention was unknown. Therefore, we judged the risk of performance bias as unclear in all studies in the review.

In four of the included studies, assessors were blinded to the treatment condition (Bieleninik 2017; Gattino 2011; Rabeyron 2020; Sharda 2018). In three further studies (Buday 1995; Lim 2010; Lim 2011), assessors were blinded to the purpose of the research. We judged all these seven studies as being at low risk of bias. In Kim 2008, non-generalised outcome measures and two of the measures assessing generalised skills (ESCS, MPIP) were rated by the researcher and complemented by independent coders (inter-rater reliability ranging from 0.70 to 0.98). We judged the risk of bias as being unclear. Studies primarily using parent reports as outcomes were judged as being at unclear risk of bias (Ghasemtabar 2015; LaGasse 2014; Yurteri 2019). In Thompson 2014, measures were based on parent reports; however, they contained internal safeguards to address bias as evidenced by high correlations with non-parent-rated measures and high test-retest correlations (e.g. Pearson's r = 0.70, P = 0.01, for the SRS's one-month test-retest reliability). Nevertheless, we judged this study to be at unclear risk. We also judged studies as being at unclear risk of bias where detailed information about assessor blinding was missing (Arezina 2011; Brownell 2002; Chen 2010; Chen 2013; Huang 2015; Mateos-Moreno 2013; Thomas 2003). In four studies, outcomes were assessed by the participants through self-report questionnaires (Moon 2010; Porter 2017; Schwartzberg 2013; Schwartzberg 2016). In two studies, outcome assessors were not blinded (Bharathi 2019; Farmer 2003), and in one study, Sa 2020, the outcome measure was applied by the therapist administering the intervention (thus yielding the data ineligible for our meta-analysis). For these remaining seven studies, we judged the risk of bias to be high.

Six studies used more than one rater to independently assess outcomes. Five of those studies reported a high inter-rater reliability for the assessment of outcomes (Arezina 2011: inter-observer agreement ranging from 85.7% to 98.9%; Brownell 2002: inter-rater reliability 0.86 to 0.94; Buday 1995: agreement rate 98%; Farmer 2003: agreement rate 91%; Kim 2008: inter-rater reliability 0.70 to 0.98, as reported above). In Mateos-Moreno 2013, measures were taken independently by two assessors. Any possible disagreement was discussed until agreement was reached on a final score to be used for analysis.

Incomplete outcome data

Twenty-three studies reported no or low attrition rates, leading to a low risk of bias judgement. Out of these studies, very low to acceptable dropout rates, ranging from 2% to 28% until the post-intervention assessment, were reported for five studies (Bieleninik 2017; Porter 2017; Rabeyron 2020; Sharda 2018; Thompson 2014). LaGasse 2014 excluded a participant with available data from the published analysis; however, the IPD-based analyses presented here included all participants. Three studies had dropout rates above 30% and were judged as entailing a high risk of bias due to attrition (Kim 2008: 5/15, 33%; Schwartzberg 2013: 77/107, 72%; Schwartzberg 2016: 64/93, 69%).

Selective reporting

There was no evidence of selective reporting of outcomes in the included studies, leading to a low risk of bias judgement.

Other potential sources of bias

We considered inadequate music therapy methods and inadequate music therapy training of therapists as additional potential sources

of bias. For the majority of studies, we detected none of these sources of bias. For Chen 2013 and Huang 2015, it was unclear whether the music therapy was provided by a trained music therapist. Moon 2010 described a music drama approach which might closely link to music therapy, but it was unclear whether this approach was provided by a trained music therapist. For Buday 1995, we found both the music therapy methods and the training of the person delivering the intervention to be of unclear adequacy.

Effects of interventions

See: **Summary of findings 1** Music therapy compared with placebo therapy or standard care for autistic people

Twenty-five of the included studies were included in the meta-analyses; in one study (Sa 2020), outcomes were measured by the therapist and therefore not eligible to be included. We used fixed-effect analyses for all outcomes, but changed to a random-effects model when a substantial amount of heterogeneity (i.e. 50% or higher; Deeks 2021) was identified immediately post-intervention that could not be explained by clinical subgroups.

Primary outcomes

Global improvement

Post-intervention

In eight studies, global improvement was assessed immediately post-intervention (Bharathi 2019; Bieleninik 2017; Kim 2008; LaGasse 2014; Porter 2017; Rabeyron 2020; Schwartzberg 2013; Thompson 2014). The RR for global improvement between music therapy and comparison groups was 1.22 (95% confidence interval (CI) 1.06 to 1.40, P = 0.006; number needed to treat for an additional beneficial outcome NNTB = 11 for low-risk population, 95% CI 6 to 39; NNTB = 6 for high-risk population, 95% CI 3 to 21; 8 studies, 583 participants; moderate-certainty evidence; Analysis 1.1), suggesting that global improvement is more likely to occur with music therapy than with 'placebo therapy' or standard care alone. There was no heterogeneity ($Chi^2 = 5.53$, P = 0.60, $I^2 = 0\%$) and therefore we did not examine potential moderators; we retained the fixed-effect model for this outcome. Changing to a random-effects model yielded similar results (P = 0.002). In a sensitivity analysis excluding data from two high-attrition studies (Kim 2008; Schwartzberg 2013), the effect for global improvement showed no substantial changes (P = 0.004).

One to five months follow-up

Two studies (Bharathi 2019; Porter 2017) also evaluated global improvement in the period one to five months post-intervention. The RR for global improvement between music therapy and comparison groups for this period was 1.19 (95% CI 0.90 to 1.57; P = 0.22; 2 studies, 99 participants), indicating no clear evidence of a difference between music therapy and comparison groups.

Six to 11 months follow-up

One study (Bieleninik 2017) measured global improvement in the period six to 11 months post-intervention. The RR for global improvement between music therapy and comparison groups for this period was 1.14 (95% CI 0.91 to 1.41, P = 0.25; 1 study, 364 participants), again indicating no clear evidence of a difference between the groups.

Social interaction

Post-intervention

Immediately post-intervention, average endpoint scores of social interaction were available from 12 studies (Bharathi 2019; Bieleninik 2017; Chen 2013; Gattino 2011; Ghasemtabar 2015; Kim 2008; LaGasse 2014; Porter 2017; Rabeyron 2020; Schwartzberg 2013; Sharda 2018; Thompson 2014). As heterogeneity was substantial ($Chi^2 = 29.51$, P = 0.002, $I^2 = 63\%$) and could not be explained clinically via subgroup analyses (results not shown), we accordingly conducted a random-effects analysis for this outcome. The SMD effect estimate was in the small to medium range, but the CI ranged from no effect to a medium effect (SMD 0.26, 95% CI −0.05 to 0.57, P = 0.11; 12 studies, 603 participants; low-certainty evidence; Analysis 1.2), thus indicating no clear evidence of a difference between music therapy and comparison groups. Investigating the related funnel plot did not yield any asymmetry, thus there was no clear indication of a risk of non-reporting bias.

During intervention

Average endpoint scores of social interaction during the intervention were available from three studies (Arezina 2011; Kim 2008; Thomas 2003) and showed a large effect (SMD 1.15, 95% CI 0.49 to 1.80, P < 0.001; 3 studies, 44 participants; Analysis 1.2), favouring music therapy over comparison groups. The results were homogeneous ($Chi^2 = 1.50$, P = 0.47, $I^2 = 0\%$). We conducted a sensitivity analysis excluding data from the high-attrition study (Kim 2008), and found that the effect for social interaction remained statistically significant (P = 0.05). No heterogeneity was detected for this analysis ($Chi^2 = 0.64$, P = 0.42, $I^2 = 0\%$).

One to five months follow-up

Effect estimates in the period one to five months post-intervention (SMD 0.54, 95% CI −0.11 to 1.19, P = 0.10; 2 studies, 59 participants) showed little to no difference between the conditions.

Six to 11 months follow-up

Effect estimates in the period six to 11 months post-intervention (SMD −0.06, 95% CI −0.30 to 0.18, P = 0.63; 1 study, 258 participants) indicated no clear evidence of a difference between music therapy and comparison groups.

Non-verbal communication

Post-intervention

Seven studies assessed non-verbal communication immediately post-intervention (Chen 2013; Gattino 2011; Kim 2008; LaGasse 2014; Rabeyron 2020; Sharda 2018; Thompson 2014). The heterogeneity found for this comparison was only moderate ($Chi^2 = 9.81$, P = 0.13, $I^2 = 39\%$), and therefore we did not examine potential moderators; we kept a fixed-effect SMD model for this outcome. The effect size for difference between music therapy and control was in the small to medium range, but the CI ranged from no effect to a medium effect (SMD 0.26, 95% CI −0.03 to 0.55, P = 0.08; 7 studies, 192 participants; low-certainty evidence; Analysis 1.3), suggesting little to no difference between the conditions. Changing to a random-effects model yielded similar results (P = 0.14). A sensitivity analysis excluding the study with a high dropout rate (Kim 2008) also did not lead to substantial changes in the results for generalised non-verbal communication (P = 0.15).

During intervention

Average endpoint scores for non-verbal communication during the intervention were available from three studies (Buday 1995; Farmer 2003; Kim 2008) and indicated a large effect favouring music therapy (SMD 1.06, 95% CI 0.44 to 1.69, P < 0.001; 3 studies, 50 participants; Analysis 1.3). The results showed heterogeneity (Chi^2 = 4.71, P = 0.09, I^2 = 58%), which may be related to the relatively high attrition rate in Kim 2008, or the unclear quality of music therapy methods and therapist's training in Buday 1995. When excluding data from both studies, the overall effect did not show substantial changes (SMD 1.64, 95% CI 0.10 to 3.19, P = 0.04).

Verbal communication

Post-intervention

Eight studies assessed verbal communication immediately post-intervention (Chen 2013; Gattino 2011; Lim 2010; Lim 2011; Rabeyron 2020; Schwartzberg 2013; Sharda 2018; Thompson 2014). The results showed substantial heterogeneity (Chi^2 = 25.30, P < 0.001, I^2 = 72%) that could not be explained clinically via subgroup analyses (results not shown), resulting in a random-effects model being used for this outcome. The effect size for difference in verbal communication immediately post-intervention was in the small to medium range, but the CI ranged from no effect to a medium effect (SMD 0.30, 95% CI −0.18 to 0.78; P = 0.21; 8 studies, 276 participants; very low-certainty evidence; Analysis 1.4), suggesting little to no difference between the conditions.

During intervention

Four studies investigated verbal communication during the intervention (Buday 1995; Farmer 2003; Schwartzberg 2013; Schwartzberg 2016). The CI of the effect estimate for difference in verbal communication during the intervention ranged from a medium harmful effect to a small to medium beneficial effect (SMD −0.06, 95 % CI −0.41 to 0.28, P = 0.71; 4 studies, 129 participants; Analysis 1.4), indicating no clear evidence of a difference between the groups. There was no heterogeneity (Chi^2 = 1.95, P = 0.58, I^2 = 0%).

One to five months follow-up

Data for verbal communication, measured in the period of one to five months post-intervention using a standardised scale, were available from one study (Bharathi 2019). The SMD effect size for this follow-up period was small, but the CI ranged from a small to medium harmful to a large beneficial effect (SMD 0.22, 95% CI −0.33 to 0.76, P = 0.44; 1 study, 52 participants; Analysis 1.4), indicating no clear evidence of a difference between music therapy and comparison groups, a similar finding to the other time points for this outcome.

Quality of life

Post-intervention

Three studies investigated quality of life (QoL) of participants and/or their families immediately post-intervention (Bieleninik 2017; Sharda 2018; Yurteri 2019). The SMD effect size across studies was 0.28 (95% CI 0.06 to 0.49, P = 0.01; 3 studies, 340 participants; moderate-certainty evidence; Analysis 1.5), indicating a small to medium effect favouring music therapy, which suggests that music therapy probably increases QoL compared with 'placebo therapy' or standard care alone. Heterogeneity was low (Chi^2 = 2.41, P = 0.30, I^2 = 17%) and therefore we did not examine potential moderators; we retained the fixed-effects model for this outcome. Changing to a random-effects model did not lead to substantial changes of the results (P = 0.02).

Six to 11 months follow-up

One of the studies also measured QoL seven months after the end of the intervention, i.e. in the period six to 11 months post-intervention (Bieleninik 2017). The CI of the effect estimate for difference in quality of life in this period ranged from a small harmful to a small to medium beneficial effect (SMD 0.04, 95% CI −0.21 to 0.29, P = 0.73; 1 study, 249 participants), indicating no clear evidence of a difference between music therapy and comparison groups.

Total autism symptom severity

Post-intervention

Nine studies assessed total autism symptom severity immediately post-intervention (Bharathi 2019; Bieleninik 2017: Chen 2010; Chen 2013; Huang 2015; LaGasse 2014; Mateos-Moreno 2013; Rabeyron 2020; Yurteri 2019). The results showed substantial heterogeneity (Chi^2 = 61.33, P < 0.001, I^2 = 87%) that could not be explained clinically via subgroup analyses (results not shown), so we chose a random-effects model. The effect size for difference in total autism symptom severity immediately post-intervention was large (SMD −0.83, 95% CI −1.41 to −0.24, P = 0.005; 9 studies, 575 participants; moderate-certainty evidence; Analysis 1.6), suggesting that music therapy probably decreases total autism symptom severity compared to 'placebo therapy' or standard care alone.

During the intervention

Total autism symptom severity during the intervention was measured in one study (Mateos-Moreno 2013). The effect estimate was small, with a wide CI (SMD 0.15, 95% CI -0.83 to 1.14, P = 0.76; 1 study, 16 participants; Analysis 1.6), indicating no clear evidence of a difference between music therapy and comparison groups.

One to five months follow-up

Average endpoint scores of total autism symptom severity measured in the period of one to five months post-intervention were available from two studies (Bharathi 2019; LaGasse 2014) and showed a large effect in favour of music therapy (SMD −0.93, 95% CI −1.81 to −0.06, P = 0.04; 2 studies, 69 participants).

Six to 11 months follow-up

One study also assessed total autism symptom severity seven months after the end of the intervention, i.e. in the period six to 11 months post-intervention (Bieleninik 2017). The SMD effect size for this time point was small, but the CI ranged from no effect to a small to medium effect (SMD 0.18, 95% CI −0.05 to 0.41, P = 0.13; 1 study, 289 participants), indicating no certain differences between music therapy and comparison groups.

Adverse events

Data for adverse events immediately post-intervention and in the period six to 11 months post-intervention were available from two studies (Bieleninik 2017; Porter 2017). However, as no events occurred in Porter 2017, only Bieleninik 2017 contributed an effect estimate (Analysis 1.7). Adverse events were rare,

and no differences were observed between music therapy or standard care in either time period (RR 1.52, 95% CI 0.39 to 5.94, P = 0.55 immediately post-intervention, 1 study, 290 participants; RR 0.88, 95% CI 0.23 to 3.46, P = 0.86 at 6-11 months post-intervention, 1 RCT, 290 participants; moderate-certainty evidence), indicating similar frequencies of adverse events in both trial arms. Bieleninik 2017 reported that adverse events included hospitalisation and institutional stay, as reported by parents, and mainly referred to planned and short-term institutional stays. Porter 2017 reported that no serious adverse events or non-serious adverse events attributable to either arm of the trial occurred (personal communication, 25 January 2021). No other adverse events were reported in any of the other included studies.

Secondary outcomes

Adaptive behaviour

Post-intervention

Immediately post-intervention, average endpoint scores of adaptive behaviour were available from five studies (Bieleninik 2017; Chen 2010; Porter 2017; Rabeyron 2020; Sharda 2018). The CI of the effect estimate for difference in adaptive behaviour at this time point ranged from no effect to a small effect (SMD −0.02, 95% CI −0.20 to 0.16, P = 0.84; 5 studies, 462 participants; Analysis 1.8), indicating no differences between music therapy and comparison groups. No heterogeneity was detected for this comparison (Chi^2 = 0.62, P = 0.96, I^2 = 0%), so we did not examine potential moderators and retained the fixed-effects model for this outcome. Changing to a random-effects model did not lead to substantial changes of the results (P = 0.84).

During the intervention

Four studies investigated adaptive behaviour during the intervention (Arezina 2011; Brownell 2002; Kim 2008; Thomas 2003). The SMD effect size for difference between music therapy and 'placebo' therapy groups was 1.19 (95 % CI 0.56 to 1.82, P < 0.001; 4 studies, 52 participants; Analysis 1.8), indicating a large effect in favour of music therapy. Heterogeneity was low (Chi^2 = 4.16, P = 0.24, I^2 = 28%). The effect on adaptive behaviour during the intervention remained large and homogeneous in a sensitivity analysis excluding two studies with high risk of bias (Brownell 2002; Kim 2008).

One to five months follow-up

Effects in the period one to five months post-intervention (Porter 2017; SMD 0.56, 95% CI −0.12 to 1.24, P = 0.11; 1 study, 35 participants) indicated no clear evidence of a difference between music therapy and comparison groups.

Six to 11 months follow-up

Similarly, effects in the period six to 11 months post-intervention (Bieleninik 2017; SMD −0.12, 95% CI −0.36 to 0.11, P = 0.29; 1 study, 290 participants) indicated no clear evidence of a difference between music therapy and comparison groups.

Quality of family relationships

Post-intervention

Three studies assessed the quality of family relationships immediately following the intervention (Kim 2008; Porter 2017; Thompson 2014). The effect size for difference between music therapy and control groups was in the small to medium range, but the CI ranged from a small harmful to a large beneficial effect (SMD 0.29, 95% CI −0.24 to 0.83, P = 0.28; 3 studies, 56 participants; Analysis 1.9), indicating no clear evidence of a difference between music therapy and comparison groups. There was no indication of heterogeneity between studies (Chi^2 = 0.37, P = 0.83, I^2 = 0%), therefore we did not examine potential moderators and retained the fixed-effects model for this outcome. Changing to a random-effects model yielded similar results (P = 0.28).

One to five months follow-up

For follow-up in the period of one to five months post-intervention, the CI of the effect estimate ranging from a large harmful to a large beneficial effect indicated uncertain differences between music therapy and standard care (Porter 2017; SMD −0.04, 95% CI −1.07 to 0.99, P = 0.94; 1 study, 15 participants).

Identity formation

Post-intervention

Two studies assessed aspects of identity formation (including self-esteem, self-awareness, and self-efficacy) immediately post-intervention (Moon 2010; Porter 2017). The results showed substantial heterogeneity (Chi^2 = 7.82, P = 0.005, I^2 = 87%) that could not be explained clinically via subgroup analyses (results not shown), so we used a random-effects model. The SMD effect size for difference in identity formation immediately post-intervention was large, but the CI ranged from a medium harmful to a large beneficial effect (SMD 1.35, 95% CI −0.58 to 3.28, P = 0.17; 2 studies, 55 participants; Analysis 1.10), indicating no clear evidence of a difference between music therapy and comparison groups.

One to five months follow-up

For the period of one to five months post-intervention, results from one study (Porter 2017) for self-esteem indicated a large effect in favour of music therapy (SMD 0.86, 95% CI 0.16 to 1.55, P = 0.02; 1 study, 35 participants).

Depression

Post-intervention

Depression was assessed in one study (Porter 2017). Results showed no clear evidence of a difference between music therapy and treatment-as-usual (SMD −0.34, 95% CI −1.01 to 0.34, P = 0.33; 1 study, 34 participants; Analysis 1.11).

One to five months follow-up

There was little to no difference between the conditions at the period of one to five months post-intervention (SMD −0.60, 95% CI −1.27 to 0.07, P = 0.08; 1 study, 36 participants).

Cognitive ability

Post-intervention

One study assessed aspects of cognitive ability by measuring attention skills immediately post-intervention (Sa 2020). However, data from this study were not included as the outcome measure was applied by the therapist.

DISCUSSION

Summary of main results

We found 26 trials that evaluated the effects of music therapy for autistic individuals aged two years to young adult age. Outcomes were assessed during the intervention, immediately post-intervention, and within two periods of follow-up post-intervention (one to five months; six to 11 months after the end of therapy). Music therapy was compared with standard care, or with a 'placebo' therapy which attempted to control for all non-specific elements of music therapy, such as the use of music or the attention of a therapist.

The results show evidence of a large effect in favour of music therapy on social interaction during the intervention. However, the certainty of the evidence using the GRADE system (Schünemann 2013) was rated as 'low' meaning that our confidence in the effect estimate is limited. There was also a large effect in favour of music therapy on non-verbal communication during the intervention, again with a 'low' certainty of the evidence meaning that results should be considered with caution. In addition, a large effect in favour of music therapy was found for total autism symptom severity both immediately and one to five months post-intervention; we rated the certainty as 'moderate' for this outcome, meaning that we are moderately confident in the effect estimate.

Large effects in favour of music therapy were also found for the secondary outcomes adaptive behaviour during the intervention, and for identity formation one to five months post-intervention. Small to moderate effect sizes resulted for the primary outcomes global improvement and quality of life immediately post-intervention. The certainty of the evidence was rated as 'moderate' for these two outcomes, meaning that the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

No evidence of effect was found for the primary outcome verbal communication (rated as 'very low' certainty of the evidence immediately post-intervention), and for secondary outcomes quality of family relationships and depression. For adverse events, no differences were found between music therapy and standard care ('moderate' certainty of the evidence, which means that we are moderately confident about this result).

It is interesting to note that social interaction and non-verbal communication skills, which may be more closely related to non-verbal interaction occurring within music therapy, showed change compared with no change detected for verbal communication. However, it may also be that social interaction and non-verbal communication skills are relatively easier to address than verbal communication skills, especially in minimally verbal children and through short- to medium-term interventions. It is also interesting to see that both social interaction and non-verbal communication showed change during the intervention, but not following the intervention. Generalising skills acquired within the intervention context to novel contexts and across interaction partners is a known difficulty in autism (Green 2018), and it may be that helping individuals in generalising their skills requires longer periods of intervention, and/or different approaches such as including the individual's everyday/family system in interventions. Considering this challenge of skill generalisation across contexts, it is remarkable to see the benefits of music therapy based on measures outside of the therapy environment and after completion of therapy in four outcome areas (global improvement, quality of life, total autism symptom severity, and identity formation).

Overall completeness and applicability of evidence

Music therapy conditions

An important improvement regarding the applicability of the evidence in this update is that it includes more clinical techniques and components of music therapy that are in line with those used in clinical practice. The early studies that were included in the first version of this review (Gold 2006) were of limited generalisability to clinical practice (Brownell 2002; Buday 1995; Farmer 2003). These studies only used a limited subset of the music therapy techniques described in the clinical literature in the experimental treatment conditions. Receptive music therapy techniques with a high level of structuring predominated in those interventions; improvisational techniques were not utilised. However, active and improvisational techniques are widely used in many parts of the world (Edgerton 1994; Gattino 2011; Geretsegger 2015; Holck 2004; Kim 2006; Schumacher 1999a; Schumacher 1999b; Thompson 2014; Thompson 2012; Wigram 2006; Wigram 2009). In addition to five of the seven studies added in the previous review update (Geretsegger 2014), 20 newly included studies added in this review update reflect active techniques. Most of them emphasise relational and/or improvisational approaches to music therapy, thus considerably increasing the applicability of findings to clinical practice and hence the external validity of this review.

In terms of therapy setting, it is noteworthy that in about half of the newly added studies in this update, music therapy was provided in a group setting, while group music therapy was not applied in any of the studies in the previous version of this review (Geretsegger 2014; except for the family-based setting in Thompson 2014). We did not plan to analyse the comparative effects of these settings. Clinically, both settings may be relevant for different individuals, or for the same individuals at different times. Although a group setting may be overwhelming for some autistic individuals, music therapy groups with a tailored mix of structured and more flexible elements may also provide valuable opportunities for engaging in predictable and pleasurable social interactions with a variety of persons. This aspect of our review may also help in applying this review's findings to contemporary healthcare settings where individual therapy is often difficult to obtain due to economical reasons.

Generally speaking, music therapy for autistic individuals should be backed by research evidence from both music therapy and related fields, aiming at cooperation with others involved in treatment and care of clients, active engagement of clients, and establishing structure, predictability, and routines. It is important to note that providing structure does not equal rigidity within interventions. Music contains rhythmic, melodic, harmonic, and dynamic structure which, when applied systematically and skilfully, can be effective in engaging autistic children. Intervention strategies employing music improvisation are usually not pre-structured in the sense of a fixed manual. In recent years, flexible but systematic treatment protocols for music therapy have been developed in clinical practice and research investigations in autism (Geretsegger 2015; Kim 2006; Thompson 2014; Wigram 2006) as well as in other fields (Baker 2019; Millstein 2021; Rolvsjord 2005). As

described above (see Included studies), several of the studies in this review have successfully applied such guidelines. More studies employing therapy approaches which are close to those applied in clinical practice will be needed in order to further improve the clinical applicability of research findings.

Control conditions

Thirteen of the included studies used a 'dismantling' strategy to isolate the effect of the specific 'ingredients' of music therapy by setting up comparison conditions which were very similar to the music therapy interventions, excluding only the music component (Arezina 2011; Brownell 2002; Buday 1995; Farmer 2003; Kim 2008; LaGasse 2014; Lim 2010; Lim 2011; Moon 2010; Schwartzberg 2013; Schwartzberg 2016; Sharda 2018; Thomas 2003). Any conclusion from such comparisons will therefore address the effects of specific music therapy techniques, rather than the absolute effects of music therapy in general. This type of design is most justified in explanatory trials (Thorpe 2009) or when exploring music therapy intervention strategies. However, such comparison conditions are less appropriate in pragmatic trials designed to inform practice (Thorpe 2009) as they may introduce some artificiality into the studies through selecting out and applying a single intervention strategy. This is not typically undertaken in clinical treatment, although it does isolate specific components of music therapy. In the broader field of psychotherapy research, similar constructions of 'placebo' therapy to control for the therapist's attention and the non-specific elements have been broadly used (Kendall 2004, pp. 20-1). However, research on common factors in psychotherapy raise the question of how adequate it is conceptually, and also whether it is technically possible to separate the active from the non-active elements of therapy (Lambert 2004, pp. 150-2).

Duration, population, and outcomes

Autism as a pervasive developmental disorder is a chronic condition, which requires sustained therapeutic intervention starting as early as possible. In clinical reports for autism, music therapy is usually described as a longer-term intervention and, given the typical emergence of entrenched and deteriorating behaviour, therapeutic intervention relies on consolidating progress over time. With the therapy duration of included studies ranging up to eight months, we consider this review's findings as sufficiently applicable to clinical contexts.

With regards to the population addressed, it is noteworthy that, different from the previous version of this review (Geretsegger 2014) which only included studies with children up to nine years, this update included studies with adolescents and young adults. The applicability of the findings is still limited to the age groups included in the studies (two years to young adult age). No direct conclusions can be drawn about music therapy in autistic individuals above the young adult age. As with most autism research, the majority of the participants in this review were males from Western countries. It is positive that some included studies have been conducted in non-Western countries. To improve generalisability to the broader population, it will be important to further diversify the populations studied in future trials to include non-male, non-Western participants.

The outcomes addressed in the included studies cover areas that form the core of the condition and relevant related areas that we consider as highly relevant to autistic individuals and their families. Having said that, it is also important to consider possible detrimental effects of approaches aiming at reducing autism severity, particularly in the areas of social interaction and communication. Such approaches might support or even provoke the masking of autistic traits, which has been reported to be associated with negative consequences for mental health including an increase in the risk of lifetime suicidality (Cassidy 2020). Additionally, the concept of autism severity and functioning-level descriptors such as 'high-functioning' are highly contentious and have recently shown to be an imprecise understanding of autistic peoples' specific needs; it has been suggested to instead acknowledge that the level of support needs of autistic people likely varies across domains, so that describing support needs in different domains (e.g. unstructured recreation activities, academic work) would be more appropriate (Bottema-Beutel 2021). However, as described above, many of the included studies employed music therapy approaches where therapists follow the individual's strengths and resources in an effort to maximise the individual's capabilities rather than to simply decrease autism symptoms or teaching specific skills for neurotypical interaction and communication. This and an emerging focus on outcomes such as quality of life, depression, and identity formation enhance the relevance of this review's findings for autistic individuals.

Quality of the evidence

Using the GRADE system (Schünemann 2013), we rated the certainty of the evidence as 'moderate' for four outcomes, 'low' for two outcomes, and 'very low' for one outcome included in the Summary of findings 1, which means that further research is likely to change the effect estimates and our confidence that they are precise; results should therefore be considered with caution. Our assessments of the certainty of the evidence mainly reflect concerns about risk of bias and imprecision due to wide CIs and small sample sizes. Limitations to the methodological strength of the evidence are due to poor reporting of randomisation and allocation procedures or lack of randomisation and/or concealment in some studies. When interpreting the results, it is important to note that, due to the nature of the intervention, it was not possible to blind those who delivered music therapy or those who received it. However, although participants were not blinded, this was unlikely to introduce bias as they are usually not fully aware of available therapy options or study design (Cheuk 2011). Additionally, blinding of assessors was not assured in the majority of studies as some of the measures in the included studies relied on reports from parents or participants themselves who were aware of the respective group allocation. However, change in participants' skills as assessed by parents or self-report may reflect effects of interventions that are meaningful and relevant to clients and their families and is therefore considered important to include.

Overall, we also observed several positive trends in this update that improve the certainty of the evidence: Most notably, both the median number of participants per study and the total number of participants included have considerably increased (from a median of 10 and a total of 165 participants in the previous review to 24 and 1165 participants, respectively, in the present update). Studies also employed longer periods of intervention on average, and a fifth of the studies in this review also included follow-up assessments ranging from three weeks until seven months after the end of the intervention, thus providing important information regarding the question of whether the effects of music therapy are enduring. It is also noteworthy that the number of studies in this

review that used validated scales (usually measuring generalised behaviour) has substantially increased; thus, the findings are both more relevant and more reliable, and more comparable across interventions.

Potential biases in the review process

One can never be completely sure that all relevant trials have been identified. However, our searches included not only exhaustive electronic and handsearches, but relied additionally on an existing international network of leading researchers in the field. Therefore, it seems unlikely that an important trial exists that did not come to our attention. Furthermore, this field does not seem to be characterised by strongly selective publication. The trials that were unpublished or published only in the grey literature tended to have positive results and were unpublished for reasons unrelated to study results (Arezina 2011; Thomas 2003).

The potential bias regarding the inclusion of studies in which one or more review authors were involved (Bieleninik 2017; Kim 2008; Thomas 2003) was mitigated by ensuring that eligibility, risk of bias and certainty of evidence assessment and data extraction were performed by two independent reviewers not involved in these studies.

We found five ongoing studies (one of which is awaiting classification due to incomplete information regarding eligibility); incorporating these studies in a future update may alter the conclusions of this review.

Agreements and disagreements with other studies or reviews

The findings of the present systematic review add substantial and relevant information to previous works about the effectiveness of music therapy for autistic people (Gold 2006; James 2015; Marquez-Garcia 2021; Mayer-Benarous 2021; Wheeler 2008; Whipple 2004; Whipple 2012).

Focusing on the most recent reviews on the topic, Marquez-Garcia (Marquez-Garcia 2021) summarized 36 longitudinal and retrospective peer-reviewed studies published between 2008 and 2018. The review examined family interaction, communication, psychological, and physiological changes. The authors concluded that the poor methodology of the included studies (e.g. experimental designs, sample sizes, outcome measures) prevented them from recommending music therapy in this population. They also encouraged the integration of behavioural evaluations with neuroscience (e.g. neuroimaging) and a more detailed characterisation of study participants (e.g. severity level, presence of intellectual disability).

Mayer-Benarous 2021 evaluated the efficacy of educational and improvisational music therapy in children with neurodevelopmental disorders such as ASD, attention deficit-hyperactivity disorder (ADHD), and learning and intellectual disabilities. The authors principally analysed outcomes related to socio-communication. Evidence on the efficacy of educational music therapy was based on 12 studies and supported a positive but small effect of educational music therapy for autistic children. According to the nine studies evaluating improvisational music therapy, efficacy appeared limited, but promising. Similarly to Marquez-Garcia 2021, Mayer-Benarous 2021 highlighted the methodological issues of the included studies.

The findings of the present meta-analysis add considerably to the external validity of older and more recently published systematic reviews. First, the methodology was more rigorous, with clear predefined inclusion/exclusion criteria, especially concerning the population under study, the type of intervention, and the study design. Second, our systematic review was more inclusive in terms of timeframe, age of participants, and outcomes examined; of note, electronic searches were combined with a consultation of the grey literature and experts in the field. Most importantly, we performed not only a qualitative, but also a quantitative synthesis, which may allow a clearer and more objective interpretation of findings, especially in light of the scattered outcome measures adopted in the included trials. The evaluation of outcomes with immediate relevance for autistic individuals, such as quality of life, identity formation, and depression may add a considerable value to the results of the present review. Notwithstanding, in agreement with the most recent systematic reviews on the topic (Marquez-Garcia 2021; Mayer-Benarous 2021), we have underlined the urgent need for improving the methodology of trials evaluating the efficacy of music therapy for autistic people.

AUTHORS' CONCLUSIONS

Implications for practice

The evidence compiled in this review suggests that music therapy is probably associated with an increased chance of global improvement, and likely results in a small improvement in quality of life and a large improvement in total autism symptom severity immediately post-intervention. It may also improve social interaction and non-verbal communication during the intervention but not after the intervention. The evidence for verbal communication is uncertain. The evidence in our review also suggests that music therapy may improve adaptive behaviour in autistic children during the intervention but not after the intervention, and identity formation in autistic children and adolescents measured in the period of one to five months after the end of the intervention, but not immediately after the intervention. Music therapy has been shown to be superior to standard care and to similar forms of therapy where music was not used, which may be indicative of a specificity of the effect of music within music therapy.

Certain behaviours of autistic children, adolescents and adults such as self-injurious or aggressive behaviour may be a challenge to their parents and other family members (Oono 2013). Therefore, the increases in adaptive behaviour and in quality of life through music therapy as found in this review may be highly relevant findings for families affected by autism.

The possible positive effect of music therapy for social interaction and non-verbal communication measured during, but not after the intervention might be related to the known challenge of generalising skills acquired within the intervention context to novel contexts and across interaction partners. It may be conducive for skill generalisation across contexts if family members are included in therapy sessions (as done in Thompson 2014) and/or informed about and trained in relevant music-based techniques and approaches that help in creating opportunities for mutual social engagement (Gottfried 2016).

As only short- to medium-term effects up to 12 months have been examined, it remains unknown how enduring the effects of music

therapy are in the longer term. However, we found some evidence that positive effects of music therapy can be maintained after the intervention has ended. Effects on outcomes measured at follow-up in the period of one to five months post-intervention showed a possible positive effect of music therapy for total autism symptom severity and self-esteem. For other outcomes and other follow-up time points, no clear evidence of differences between music therapy and comparison groups was found.

This review suggests that music therapy probably does not increase adverse events. However, when applying the results of this review to practice, it is important to note that the application of music therapy requires academic and clinical training in music therapy. Trained music therapists and academic training courses are available in many countries, and information is usually accessible through professional associations. Training courses in music therapy teach not only the clinical music therapy techniques as described in the background of this review, but also aim at developing the therapist's personality and clinical sensitivity, which is necessary to apply music therapy responsibly.

Implications for research

The evidence included in this review centres on children and young adults, meaning that the findings are not generalisable to autistic adults. Research is needed examining effects of music therapy for autistic individuals above the young adult age.

We recommend that future trials on music therapy in this area should be: (1) pragmatic; (2) conscious of types of music therapy; (3) conscious of relevant outcome measures; and (4) include long-term follow-up assessments.

(1) Pragmatic trials of effectiveness: The earliest trials included in this review tended to be designed as efficacy or explanatory trials. Such trials are designed with internal validity in mind and are limited in their generalisability. According to Thorpe 2009, explanatory trials tend to use inflexible experimental interventions, inflexible comparison interventions, and outcomes that are not directly relevant to autistic people, but rather an indicator of a direct intervention effect. Their relevance to informing practice may be limited. Many of the more recent trials included in this review (see Included studies) have used more flexible interventions, standard care comparisons, and downstream outcomes. Further pragmatic trials should use rigorous designs in order to reliably address the question of effectiveness (i.e. whether music therapy works 'under usual conditions', Thorpe 2009). For increasing the methodological quality of trials and reducing risk of bias, standards on randomisation, allocation, and blinding procedures should be followed and reported more strictly.

(2) Types of music therapy: As discussed in this review, various types of music therapy have been proposed. Future trials should continue to be conscious of the quality, clinical applicability and link to usual practice, and type of music therapy examined, and also investigate heterogeneity in populations (i.e. what works for whom; for example, regarding levels of support, verbal skills, socioeconomic status, or cultural background). Future trials might entail comparisons between types and settings of music therapy, but should also continue to investigate music therapy compared with other interventions or standard care. As online delivery of music therapy services is currently an emerging area of practice (Gaddy 2020), it is important to note that, in the studies included in this review, this modality has not been applied. Due to the specific benefits and limitations of online service delivery, it will be important for future studies to also examine the effects of online music therapy for autistic individuals.

(3) Relevant outcome measures: There is currently no consensus about the most pertinent outcome measures to be used in autism intervention research (McConachie 2015; Provenzani 2020; Warren 2011; Wheeler 2008). However, in line with recommendation (1) above, future trials should include outcomes that address the core problems of ASD in a generalised setting utilising standardised scales. They should also apply outcomes that are regarded as important by autistic people and their family members (McConachie 2015). Participatory approaches to research that incorporate the views of autistic people and those who support them in all stages of the research process are an important avenue to ensure that research yields relevant benefits and improved outcomes for autistic people (Fletcher-Watson 2019). When viewing social interaction as a shared responsibility and participatory practice situated within a historical, cultural, and social ecology (De Jaegher 2007; Milton 2019), measuring social skills on neurotypical premises is likely to fail in capturing progress meaningfully for autistic people. Hence, future research would benefit from incorporating embodied and enactive social cognitive perspectives, taking into account the disabling impact any given interaction, context or environment can hold for autistic people, when designing studies and choosing or developing outcomes. Outcome domains outside of core symptom areas such as psychiatric disorders which are highly prevalent in autistic adults (Lipinski 2019) should also be considered, particularly as music therapy has also shown to be beneficial for mental health conditions such as depression and anxiety in neurotypical populations (Aalbers 2017). Finally, combining biological markers with behavioural measures as done in one study in this review (Sharda 2018) may yield important findings about underlying neurobiological mechanisms in music therapy for autism (Sharda 2019).

(4) Long-term follow-up assessments: Although an increasing number of studies in this update have addressed extended time periods compared with earlier studies, only one study to date has examined outcomes up to 12 months from randomisation. With the increasing prevalence of parallel trials, long-term follow-up assessments are becoming feasible and should be considered. Examples of other psychosocial interventions for autism that failed to show effects at 12 months but showed effects after five years (Pickles 2016) should be encouraging.

ACKNOWLEDGEMENTS

We thank Margaret Anderson, Information Specialist of the Cochrane Developmental, Psychosocial and Learning Problems Group (CDPLPG), for undertaking the searches for this and the previous update, and Joanne Duffield and Sarah Davies, Managing Editors of the CDPLPG, for their kind guidance and assistance throughout the review process.

We thank Xi-Jing Chen, Jinah Kim, and Eajin Yoo for their kind help in retrieving Chinese and Korean studies, respectively, and for providing support in translating them. Thanks also go to Asli Doğrul and Ayşegül Yay Pençe for translating information from Turkish studies, to Xueli Tan and Yaning Wu for translating information from

Chinese studies, and to Eun Jin Nausner for translating information from Korean studies.

We also thank all study authors who kindly provided additional data from their respective studies.

The CRG Editorial Team are grateful to the following peer reviewers for their time and comments: Dr Catherine E Carr, Unit for Social and Community Psychiatry, Queen Mary University of London, UK; Michelle Hintz, PsyD, MT-BC, Cadenza Center for Psychotherapy & the Arts, Hollywood, FL; Ming Yuan Low, PhD, MT-BC, Berklee College of Music, Boston, MA; and Vanessa M Young, MS, School of Social and Behavioral Sciences, Arizona State University.

The CRG Editorial Team are grateful to Anne Lethaby for copyediting this review.

REFERENCES

References to studies included in this review

Arezina 2011 *{published and unpublished data}*

Arezina CH.The Effect of Interactive Music Therapy on Joint Attention Skills in Preschool Children with Autism Spectrum Disorder [Master's thesis]. Lawrence (KS): University of Kansas, 2011.

Bharathi 2019 *{published and unpublished data}*

Bharathi G, Venugopal A, Vellingiri B.Music therapy as a therapeutic tool in improving the social skills of autistic children. *Egyptian Journal of Neurology, Psychiatry and Neurosurgery* 2019;**55**:44. [DOI: 10.1186/s41983-019-0091-x]

Bieleninik 2017 *{published and unpublished data}***78923965**

* Bieleninik Ł, Geretsegger M, Mössler K, Assmus J, Thompson G, Gattino G, et al.Effects of improvisational music therapy vs enhanced standard care on symptom severity among children with autism spectrum disorder: the TIME-A randomized clinical trial. *JAMA* 2017;**318**(6):525-35. [DOI: 10.1001/jama.2017.9478] [PMCID: PMC5817481] [PMID: 28787504]

Crawford MJ, Gold C, Odell-Miller H, Thana L, Faber S, Assmus J, et al.International multicentre randomised controlled trial of improvisational music therapy for children with autism spectrum disorder: TIME-A study. *Health Technology Assessment* 2017;**21**(59):1-40. [DOI: 10.3310/hta21590] [PMCID: PMC5672498] [PMID: 29061222]

Geretsegger M, Holck U, Gold C.Randomised controlled trial of improvisational music therapy's effectiveness for children with autism spectrum disorders (TIME-A): study protocol. *BMC Pediatrics* 2012;**12**:2. [DOI: 10.1186/1471-2431-12-2] [PMCID: PMC3280156] [PMID: 22221670]

ISRCTN78923965.Trial of improvisational music therapy's effectiveness for children with autism [Randomised controlled trial of improvisational music therapy's effectiveness for children with autism spectrum disorders (TIME-A)]. doi.org/10.1186/ISRCTN78923965 (first received 18 June 2011).

Brownell 2002 *{published and unpublished data}*

Brownell MD.Musically adapted social stories to modify behaviors in students with autism: four case studies. *Journal of Music Therapy* 2002;**39**(2):117-44. [DOI: 10.1093/jmt/39.2.117] [PMID: 12213082]

Buday 1995 *{published data only}*

Buday EM.The effects of signed and spoken words taught with music on sign and speech imitation by children with autism. *Journal of Music Therapy* 1995;**32**(3):189-202. [DOI: 10.1093/jmt/32.3.189]

Chen 2010 *{published data only}*

Chen L, Tao HM, Li H, Shao Z, Yao L, Liao LJ.Empirical study of clinical effect of music therapy on children with autism [音乐治疗干预自闭症儿童临床效果的实证研究]. *Chongqing Medicine* 2010;**39**(18):2481-2, 2485. [DOI: 10.3969/j.issn.1671-8348.2010.18.042]

Chen 2013 *{published data only}*

Chen L, Zhang YR, Shao Z, Zhao P, Liao LJ.Observation on the curative effect of music therapy with social stories as the main content in intervention of children with autism [以社交故事为主要内容的音乐疗法干预孤独症儿童的疗效观察]. *Chinese Journal of Physical Medicine and Rehabilitation* 2013;**35**(7):575-8. [DOI: 10.3760/cma.j.issn.0254-1424.2013.07.019]

Farmer 2003 *{published data only}*

Farmer KJ.The Effect of Music vs Nonmusic Paired with Gestures on Spontaneous Verbal and Nonverbal Communication Skills of Children with Autism Between the Ages 1-5 [Master's thesis]. Tallahassee (FL): Florida State University, 2003.

Gattino 2011 *{published data only}*

Gattino GS, Riesgo RdS, Longo D, Leite JC, Faccini LS.Effects of relational music therapy on communication of children with autism: a randomized controlled study. *Nordic Journal of Music Therapy* 2011;**20**(2):142-54. [DOI: 10.1080/08098131.2011.566933]

Ghasemtabar 2015 *{published data only}*

Ghasemtabar SN, Hosseini M, Fayyaz I, Arab S, Naghashian H, Poudineh Z.Music therapy: an effective approach in improving social skills of children with autism. *Advanced Biomedical Research* 2015;**4**:157. [DOI: 10.4103/2277-9175.161584] [PMCID: PMC4550953] [PMID: 26380242]

Huang 2015 *{published data only}*

Huang LY.Rehabilitation applications of music therapy in children with autism [音乐治疗在自闭症儿童康复中的应用]. *China Continuing Medical Education* 2015;**7**(18):153-4.

Kim 2008 *{published and unpublished data}*

Kim J, Wigram T, Gold C.Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. *Autism* 2009;**13**(4):389-409. [DOI: 10.1177/1362361309105660] [PMID: 19535468]

* Kim J, Wigram T, Gold C.The effects of improvisational music therapy on joint attention behaviors in autistic children: a randomized controlled study. *Journal of Autism and Developmental Disorders* 2008;**38**(9):1758-66. [DOI: 10.1007/s10803-008-0566-6] [PMID: 18592368]

Kim J.The Effects of Improvisational Music Therapy on Joint Attention Behaviours in Children with Autistic Spectrum Disorder [PhD thesis]. Aalborg, Denmark: Aalborg University, 2006.

LaGasse 2014 *{published and unpublished data}*

LaGasse AB.Effects of a music therapy group intervention on enhancing social skills in children with autism. *Journal of Music Therapy* 2014;**51**(3):250-75. [DOI: 10.1093/jmt/thu012] [PMID: 25053766]

Lim 2010 *{published data only}*

* Lim HA.Effect of "developmental speech and language training through music" on speech production in children with autism spectrum disorders. *Journal of Music Therapy* 2010;**47**(1):2-26. [DOI: 10.1093/jmt/47.1.2] [PMID: 20635521]

Lim HA.The Effect of "Developmental Speech-Language Training through Music" on Speech Production in Children with Autism Spectrum Disorders [PhD thesis]. Coral Gables (FL): University of Miami, 2007. [WEB PAGE: citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.5692&rep=rep1&type=pdf]

Lim 2011 *{published data only}*

Lim HA, Draper E.The effects of music therapy incorporated with applied behavior analysis verbal behavior approach for children with autism spectrum disorders. *Journal of Music Therapy* 2011;**48**(4):532-50. [DOI: 10.1093/jmt/48.4.532] [PMID: 22506303]

Mateos-Moreno 2013 *{published data only}*

Mateos-Moreno D, Atencia-Doña L.Effect of a combined dance/movement and music therapy on young adults diagnosed with severe autism. *Arts in Psychotherapy* 2013;**40**(5):465-72. [DOI: 10.1016/j.aip.2013.09.004]

Moon 2010 *{published data only}*

Moon JY.Effects of therapeutic music play based on theory of mind for children with Asperger syndrome's self-cognition, self-efficacy, and self- esteem [마음이론에 근거한 음악극이 아스퍼거 증후군 아동의자기인식과 자기효능감 및 자아존중감에 미치는 영향] . *Emotional and Behavioral Disorders Research* 2010;**26**(2):197-222. [WEB PAGE: www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE08783523]

Porter 2017 *{published and unpublished data}***96352204**

McConnell T.Data request for Cochrane review about "Music Therapy for ASD" [personal communication]. Email to: C Gold 23 November 2020.

* Porter S, McConnell T, McLaughlin K, Lynn F, Cardwell C, Braiden H-J, et al.Music therapy for children and adolescents with behavioural and emotional problems: a randomised controlled trial. *Journal of Child Psychology and Psychiatry* 2017;**58**(5):586-94. [DOI: 10.1111/jcpp.12656] [PMID: 27786359]

Porter S, Holmes V, McLaughlin K, Lynn F, Cardwell C, Braiden H-J, et al.Music in mind, a randomized controlled trial of music therapy for young people with behavioural and emotional problems: study protocol. *Journal of Advanced Nursing* 2012;**68**(10):2349-58. [DOI: 10.1111/j.1365-2648.2011.05936.x] [PMID: 22235808]

Rabeyron 2020 *{published and unpublished data}***31179114**

* Rabeyron T, Robledo Del Canto J-P, Carasco E, Bisson V, Bodeau N, Vrait F-X, et al.A randomized controlled trial of 25 sessions comparing music therapy and music listening for children with autism spectrum disorder. *Psychiatry Research* 2020;**293**:113377. [DOI: 10.1016/j.psychres.2020.113377] [PMID: 32798927]

Sa 2020 *{published data only}*

Sa V.The Effect of Music Attention Control Training (MACT) for Pre-Adolescents with Autism Spectrum Disorder [Master's thesis]. Stockton (CA): University of the Pacific, 2020.

Schwartzberg 2013 *{published and unpublished data}*

Schwartzberg ET, Silverman MJ.Effects of music-based social stories on comprehension and generalization of social skills in children with autism spectrum disorders: a randomized effectiveness study. *Arts in Psychotherapy* 2013;**40**(3):331-7. [DOI: 10.1016/j.aip.2013.06.001]

Schwartzberg 2016 *{published and unpublished data}*

Schwartzberg ET, Silverman MJ.Effects of a music-based short story on short- and long-term reading comprehension of individuals with autism spectrum disorder: a cluster randomized study. *Arts in Psychotherapy* 2016;**48**:54-61. [DOI: 10.1016/j.aip.2016.01.001]

Sharda 2018 *{published data only}***26821793**

Latif N, Di Francesco C, Custo-Blanch M, Hyde K, Sharda M, Nadig A.Joint engagement and movement: active ingredients of a music-based intervention with school-age children with autism. *NeuroRehabilitation* 2021;**48**(2):167-85. [DOI: 10.3233/NRE-208012] [PMID: 33664155]

* Sharda M, Tuerk C, Chowdhury R, Jamey K, Foster N, Custo-Blanch M, et al.Music improves social communication and auditory-motor connectivity in children with autism. *Translational Psychiatry* 2018;**8**(1):231. [DOI: 10.1038/s41398-018-0287-3] [PMCID: PMC6199253] [PMID: PMC6199253]

Thomas 2003 *{unpublished data only}*

Thomas AF, Hunter BC.The effect of music therapy on communication skills of children with autism, ages 2-3: a pilot study. In: Music Therapy: The Key to Health and Wellness. Proceedings of the Fifth American Music Therapy Association Conference; 2003 18-23 Nov; Minneapolis (MN). 2003:43.

Thompson 2014 *{published and unpublished data}*

Thompson G.Making a Connection: Randomised Controlled Trial of Family Centred Music Therapy for Young Children with Autism Spectrum Disorder [PhD thesis]. Melbourne (VIC): University of Melbourne, 2012. [URI: http://hdl.handle.net/11343/37719]

* Thompson GA, McFerran KS, Gold C.Family-centred music therapy to promote social engagement in young children with severe autism spectrum disorder: a randomized controlled study. *Child: Care, Health & Development* 2014;**40**(6):840-52. [DOI: 10.1111/cch.12121] [PMID: 24261547]

Yurteri 2019 *{published data only}*

Yurteri N, Akdemir M.The effect of music therapy on autistic symptoms and quality of life in children with autism spectrum disorder [Otizm spektrum bozukluğu olan çocuklarda müzik terapinin otizm belirtileri ve yaşam kalitesine etkisi]. *Anatolian Journal of Psychiatry [Anadolu Psikiyatri Dergisi]* 2019;**20**(4):436-41. [DOI: 10.5455/apd.12505]

References to studies excluded from this review

Bringas 2015 *{published data only}*

Bringas ML, Zaldivar M, Rojas PA, Martinez-Montes K, Chongo DM, Ortega MA, et al.Effectiveness of music therapy as an aid to neurorestoration of children with severe neurological disorders. *Frontiers in Neuroscience* 2015;**9**:427. [DOI: 10.3389/fnins.2015.00427] [PMCID: PMC4631829] [PMID: 26582974]

Cowan 2016 *{published data only}*

Cowan KB.The Use of Musically Adapted Social Stories to Build Social Skills in Individuals with Autism Spectrum Disorder [PhD thesis]. Muncie (IN): Ball State University, 2016.

Dezfoolian 2013 *{published data only}*

Dezfoolian L, Zarei M, Ashayri H, Looyeh MY.A pilot study on the effects of Orff-based therapeutic music in children with autism spectrum disorder. *Music and Medicine* 2013;**5**(3):162-8. [DOI: 10.47513/mmd.v5i3.325]

Edgerton 1994 *{published data only}*

Edgerton CL.The effect of improvisational music therapy on the communicative behaviors of autistic children. *Journal of Music Therapy* 1994;**31**(1):31-62. [DOI: 10.1093/jmt/31.1.31]

Finnigan 2010 *{published data only}*

Finnigan E, Starr E.Increasing social responsiveness in a child with autism: a comparison of music and non-music interventions. *Autism* 2010;**14**(4):321-48. [DOI: 10.1177/1362361309357747] [PMID: 20591958]

Gooding 2011 *{published data only}*

Gooding LF.The effect of a music therapy social skills training program on improving social competence in children and adolescents with social skills deficits. *Journal of Music Therapy* 2011;**48**(4):440-62. [DOI: 10.1093/jmt/48.4.440] [PMID: 22506299]

Hairston 1990 *{published data only}*

Hairston MP.Analyses of responses of mentally retarded autistic and mentally retarded nonautistic children to art therapy and music therapy. *Journal of Music Therapy* 1990;**27**(3):137-50. [DOI: 10.1093/jmt/27.3.137]

Iseri 2014 *{published data only}*

Iseri E, Guney E, Guvenc RO, Guvenc AA, Ceylan MF, Cirak E, et al.Music therapy and hormonal responses in autism. *Gazi Medical Journal* 2014;**25**(2):55-8. [DOI: 12996/gmj.2014.16]

Kern 2006 *{published data only}*

Kern P, Aldridge D.Using embedded music therapy interventions to support outdoor play of young children with autism in an inclusive community-based child care program. *Journal of Music Therapy* 2006;**43**(4):270-94. [DOI: 10.1093/jmt/43.4.270] [PMID: 17348756]

Kern 2007 *{published data only}*

Kern P, Wolery M, Aldridge D.Use of songs to promote independence in morning greeting routines for young children with autism. *Journal of Autism and Developmental Disorders* 2007;**37**(7):1264-71. [DOI: 10.1007/s10803-006-0272-1] [PMID: 17120150]

Kim 2000 *{published data only}*

Kim LH.The Effect of Rhythmic Musical Activity on Stereotypic Behavior of Children with Autistic Disorder (Master's thesis). Seoul, Korea: SookMyung Women's University, 2000.

Mendelson 2016 *{published data only}*

Mendelson J, White Y, Hans L, Adebari R, Schmid L, Riggsbee J, et al.A preliminary investigation of a specialized music therapy model for children with disabilities delivered in a classroom setting. *Autism Research and Treatment* 2016;**2016**:1284790. [DOI: 10.1155/2016/1284790] [PMCID: PMC5141315] [PMID: 27999683]

Sanglakh Goochan Atigh 2017 *{published data only}*

Sanglakh Goochan Atigh A, Akbarfahimi M, Alizadeh Zarei M.The effect of movement activities in synchronization with music on motor proficiency of children with autism. *Journal of Advanced Medical Sciences and Applied Technologies* 2017;**3**(2):61-8. [DOI: 10.18869/NRIP.JAMSAT.3.2.61]

Saperston 1973 *{published data only}*

Saperston B.The use of music in establishing communication with an autistic mentally retarded child. *Journal of Music Therapy* 1973;**10**(4):184-8. [DOI: 10.1093/jmt/10.4.184]

Thaut 1988 *{published data only}*

Thaut MH.Measuring musical responsiveness in autistic children: a comparative analysis of improvised musical tone sequences of autistic, normal, and mentally retarded individuals. *Journal of Autism and Developmental Disorders* 1988;**18**(4):561-71. [DOI: 10.1007/BF02211874] [PMID: 3215882]

Yoo 2018 *{published data only}*

Yoo GE, Kim SJ.Dyadic drum playing and social skills: implications for rhythm-mediated intervention for children with autism spectrum disorder. *Journal of Music Therapy* 2018;**55**(3):340-75. [DOI: 10.1093/jmt/thy013] [PMID: 30137544]

References to studies awaiting assessment

NCT03267095 *{published data only}*

NCT03267095.Use music therapy in autism spectrum disorder [The use of music therapy in autism spectrum disorder]. clinicaltrials.gov/ct2/show/NCT03267095 (first received 24 August 2017).

References to ongoing studies

ISRCTN18340173 *{published data only}***10.1186/ISRCTN18340173**

ISRCTN18340173.Improvisational music therapy for children with autism spectrum disorder assessed using brain imaging [Music therapy outcome study for children with Autism Spectrum Disorder through integration of child neuroimaging and neuropsychology: an exploratory study]. doi.org/10.1186/ISRCTN18340173 (first received 17 February 2019).

NCT03560297 *{published data only}*

NCT03560297.SeRenade parent-child music class program [SeRenade parent-child music class program for families of children with and without ASD]. clinicaltrials.gov/ct2/show/NCT03560297 (first received 21 May 2018).

NCT04557488 *{published data only}*

NCT04557488.Effectiveness of music therapy in social skill intervention for children with ASD/ID [Effectiveness of music therapy in social skill intervention for children with ASD/ID: a randomized controlled trial]. clinicaltrials.gov/ct2/show/NCT04557488 (first received 21 September 2020).

NCT04936048 *{published data only}*

NCT04936048.Music for autism (M4A) [Music for autism: binational randomised controlled trial of music therapy versus play therapy for autistic children]. clinicaltrials.gov/ct2/show/NCT04936048 (first received 23 June 2021).

Additional references

Aalbers 2017

Aalbers S, Fusar-Poli L, Freeman RE, Spreen M, Ket JC, Vink AC, et al.Music therapy for depression. *Cochrane Database of Systematic Reviews* 2017, Issue 11. Art. No: CD004517. [DOI: 10.1002/14651858.CD004517.pub3] [PMCID: PMC6486188] [PMID: 29144545]

Accordino 2007

Accordino R, Comer R, Heller WB.Searching for music's potential: a critical examination of research on music therapy with individuals with autism. *Research in Autism Spectrum Disorders* 2007;**1**(1):101-15. [DOI: 10.1016/j.rasd.2006.08.002]

Achenbach 2001

Achenbach TM, Rescorla LA.Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment. Burlington (VT): University of Vermont, Research Center for Children, Youth & Families, 2001.

Aldridge 2002

Aldridge D, Fachner J.Music Therapy World Info-CD ROM IV 2002. University Witten-Herdecke (accessed prior to 15 March 2022).

Alvin 1991

Alvin J, Warwick A.Music Therapy for the Autistic Child. 2nd edition. Oxford, UK: Oxford University Press, 1991.

Aman 1985

Aman MG, Singh NN, Stewart AW, Field CJ.The Aberrant Behavior Checklist: a behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency* 1985;**89**(5):485-91. [PMID: 3993694]

AMTA 1999

American Music Therapy Association.Music Therapy Research: Quantitative and Qualitative Foundations CD-ROM I. 1964-1998. American Music Therapy Association (accessed prior to 15 March 2022).

APA 2013

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 5th (DSM-5) edition. Arlington (VA): American Psychiatric Publishing, 2013.

Baker 2019

Baker FA, Tamplin J, Clark IN, Lee YC, Geretsegger M, Gold C.Treatment fidelity in a music therapy multi-site cluster randomized controlled trial for people living with dementia: the MIDDEL project intervention fidelity protocol. *Journal of Music Therapy* 2019;**56**(2):125–48. [DOI: 10.1093/jmt/thy023] [PMID: 30721998]

Ball 2004

Ball CM.Music Therapy for Children with Autistic Spectrum Disorder. London, UK: Bazian Ltd, 2004.

Bandura 1978

Bandura A.Reflections on self-efficacy. *Advances in Behaviour Research and Therapy* 1978;**1**(4):237-69. [DOI: 10.1016/0146-6402(78)90012-7]

Barthélémy 2003

Barthélémy C, Lelord G.ECAR: Échelle D'évaluation des Comportements Autistiques. Manuel (version revise) [ECAR: Autistic Behavior Rating Scale. Manual (Revised version)]. Paris, FR: Editions et Applications Psychologiques (EAP), 2003.

Bellini 2007

Bellini S, Hopf A.The development of the Autism Social Skills Profile: a preliminary analysis of psychometric properties. *Focus on Autism and Other Developmental Disabilities* 2007;**22**(2):80-7. [DOI: 10.1177/10883576070220020801]

Berger 2002

Berger D.Music Therapy, Sensory Integration and the Autistic Child. London: Jessica Kingsley Publishers, 2002.

Bergmann 2016

Bergmann T.Music therapy for people with autism spectrum disorder. In: EdwardsJ, editors(s). The Oxford Handbook of Music Therapy. Oxford Library of Psychology edition. Oxford (UK): Oxford University Press, 2016:186-209. [DOI: 10.1093/oxfordhb/9780199639755.013.35]

Bishop 1998

Bishop DV.Development of the Children's Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry* 1998;**39**(6):879-91. [PMID: 9758196]

Boso 2007

Boso M, Emanuele E, Minazzi V, Abbamonte M, Politi P.Effect of long-term interactive music therapy on behavior profile and musical skills in young adults with severe autism. *Journal of Alternative and Complementary Medicine* 2007;**13**(7):709-12. [DOI: 10.1089/acm.2006.6334] [PMID: 17931062]

Bottema-Beutel 2021

Bottema-Beutel K, Kapp SK, Lester JN, Sasson NJ, Hand BN.Avoiding ableist language: suggestions for autism

researchers. *Autism in Adulthood* 2021;**3**(1):18-29. [DOI: 10.1089/aut.2020.0014]

Brondino 2015

Brondino N, Fusar-Poli L, Rocchetti M, Provenzani U, Barale F, Politi P.Complementary and alternative therapies for autism spectrum disorder. *Evidence-Based Complementary and Alternative Medicine* 2015;**2015**:258589. [DOI: 10.1155/2015/258589] [PMCID: PMC4439475] [PMID: 26064157]

Bruscia 1998

Bruscia KE.Defining Music Therapy. 2nd edition. Gilsum (NH): Barcelona Publishers, 1998.

Bury 2020

Bury SM, Jellett R, Spoor JR, Hedley D."It defines who I am" or "It's something I have": what language do [autistic] Australian adults [on the autism spectrum] prefer? Journal of Autism and Developmental Disorders 2020 Feb 28 [Epub ahead of print]. [DOI: 10.1007/s10803-020-04425-3] [PMID: 32112234]

Busner 2007

Busner J, Targum SD.The Clinical Global Impressions scale: applying a research tool in clinical practice. *Psychiatry* 2007;**4**(7):28-37. [PMCID: PMC2880930] [PMID: 20526405]

Carpente 2009

Carpente JA.Contributions of Nordoff-Robbins Music Therapy Within a Developmental, Individual-Differences, Relationship-Based (DIR®)/Floortime™ Framework to the Treatment of Children with Autism: Four Case Studies [Dissertation]. Temple University, 2009. [DOI: http://dx.doi.org/10.34944/dspace/891]

Cassidy 2020

Cassidy SA, Gould K, Townsend E, Pelton M, Robertson AE, Rodgers J.Is camouflaging autistic traits associated with suicidal thoughts and behaviours? Expanding the interpersonal psychological theory of suicide in an undergraduate student sample. *Journal of Autism and Developmental Disorders* 2020;**50**:3638–48. [DOI: 10.1007/s10803-019-04323-3]

Chen 2020

Chen B, Linke A, Olson L, Ibarra C, Reynolds S, Müller RA, et al.Greater functional connectivity between sensory networks is related to symptom severity in toddlers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry* 2020;**62**(2):160-70. [DOI: doi:10.1111/jcpp.13268]

Cheuk 2011

Cheuk DK, Wong V, Chen WX.Acupuncture for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews* 2011, Issue 9. Art. No: CD007849. [DOI: 10.1002/14651858.CD007849.pub2] [PMID: 21901712]

Cho 2004

Cho, Y.The Effects of Personal Growth Group Counseling on Children's Self-Esteem and Improvement of Learning Motivation [Masters Thesis]. Seoul: Seoul National University of Education, 2004.

Cohen 1988

Cohen J.Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale (NJ): Lawrence Erlbaum Associates, 1988.

Cohen 1999

Cohen IL, Sudhalter V.Pervasive Developmental Disorder Behavior Inventory (PDDBI-C). New York (NY): NYS Institute for Basic Research in Developmental Disabilities, 1999.

Constantino 2005

Constantino JN, Gruber CP.Social Responsiveness Scale. Los Angeles (CA): Western Psychological Services, 2005.

Constantino 2012

Constantino JN, Gruber CP.Social Responsiveness Scale: SRS-2. Torrance (CA): Western Psychological Services, 2012.

Deeks 2021

Deeks JJ, Higgins JP, Altman DG (editors).Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

De Jaegher 2007

De Jaegher H, Di Paolo E.Participatory sense-making: an enactive approach to social cognition. *Phenomenology and the Cognitive Sciences* 2007;**6**(4):485-507. [DOI: 10.1007/s11097-007-9076-9]

De Jaegher 2013

De Jaegher H.Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience* 2013;**7**:15. [DOI: 10.3389/fnint.2013.00015] [PMCID: PMC3607806] [PMID: 23532205]

De Vries 2015

De Vries D, Beck T, Stacey B, Winslow K, Meines K.Music as a therapeutic intervention with autism: a systematic review of the literature. *Therapeutic Recreation Journal* 2015;**49**(3):220-37.

Dunn 1981

Dunn LM, Dunn LM.Peabody Picture Vocabulary Test-Revised (PPVT-R). Circle Pines (MN): American Guidance Service, 1981.

Elbourne 2002

Elbourne DR, Altman DA, Higgins JP, Curtin F, Worthington HV, Vail A.Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140-9. [DOI: 10.1093/ije/31.1.140] [PMID: 10.1093/ije/31.1.140]

Elsabbagh 2012

Elsabbagh M, Divan G, Koh YJ, Kim YS, Kauchali S, Marcín C, et al.Global prevalence of autism and other pervasive developmental disorders. *Autism Research* 2012;**5**(3):160-79. [PMID: 22495912]

Epstein 1983

Epstein NB, Baldwin LM, Bishop DS.The McMaster Family Assessment Device. *Journal of Marital and Family Therapy* 1983;**9**(2):171-80. [DOI: 10.1111/j.1752-0606.1983.tb01497.x]

Eun 2000

Eun H.The Self-Regulation Group Counseling to Improve Interpersonal Competence and its Effect on Adolescents' Self-Awareness, Other-Awareness, Interpersonal Skills and Satisfaction of Interpersonal Relationship [Doctoral dissertation]. Sungkyunkwan University, 2000.

Faulstich 1986

Faulstich ME, Carey MP, Ruggiero L, Enyart P, Gresham F.Assessment of depression in childhood and adolescence: an evaluation of the Center for Epidemiological Studies Depression Scale for Children (CES-DC). *American Journal of Psychiatry* 1986;**143**(8):1024-7. [DOI: 10.1176/ajp.143.8.1024] [PMID: 3728717]

Fenigstein 1979

Fenigstein A.Self consciousness, self attention and social interaction. *Journal of Personality and Social Psychology* 1979;**37**(1):75-86. [DOI: 10.1037/0022-3514.37.1.75]

Fenson 2007

Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick JS, Bates E.MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual. 2nd edition. Baltimore (MD): PH Brookes, 2007.

First 2004

First MB, Gibbon M.The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In: HersenM, HilsenrothMJ, Segal DL, editors(s). Comprehensive Handbook of Psychological Assessment. Vol. **2. Personality Assessment**. Hoboken (NJ): John Wiley & Sons, 2004:134-43.

Fletcher-Watson 2019

Fletcher-Watson S, Adams J, Brook K, Charman T, Crane L, Cusack J, et al.Making the future together: shaping autism research through meaningful participation. *Autism* 2019;**23**(4):943-53. [DOI: 10.1177/1362361318786721] [PMCID: PMC6512245] [PMID: 30095277]

Fusar-Poli in press

Fusar-Poli L, Thompson G, Lense MD, Gold C.Music therapy and other music-based interventions. In: MatsonJL, SturmeyP, editors(s). Handbook of Autism and Pervasive Developmental Disorder: Assessment, Diagnosis, and Treatment. New York (NY): Springer Cham, (In press).

Gaddy 2020

Gaddy S, Gallardo R, McCluskey S, Moore L, Peuser A, Rotert R, et al.COVID-19 and music therapists' employment, service delivery, perceived stress, and hope: a descriptive study. *Music Therapy Perspectives* 2020;**38**(2):157-66. [DOI: doi.org/10.1093/mtp/miaa018]

Gerard 1994

Gerard AB.Parent-Child Relationship Inventory (PCRI) Manual. Los Angeles (CA): Western Psychological Services, 1994.

Geretsegger 2015

Geretsegger M, Holck U, Carpente JA, Elefant C, Kim J, Gold C.Common characteristics of improvisational approaches in music therapy for children with autism spectrum disorder: developing treatment guidelines. *Journal of Music Therapy* 2015;**52**(2):258–81. [DOI: 10.1093/jmt/thv005] [PMID: 26019303]

Gernsbacher 2017

Gernsbacher MA.Editorial Perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry* 2017;**58**(7):859-61. [DOI: 10.1111/jcpp.12706]

Gold 2004

Gold C.The use of effect sizes in music therapy research. *Music Therapy Perspectives* 2004;**22**(2):91-5. [DOI: 10.1093/mtp/22.2.91]

Gottfried 2016

Gottfried T.Creating Bridges: Music-Oriented Counseling for Parents of Children with Autism Spectrum Disorder [PhD thesis]. Aalborg, DK: Aalborg Universitet, 2016. [DOI: 10.5278/vbn.phd.hum.00042]

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime) GRADEpro GDT: GRADEpro Guideline Development Tool.Version accessed 4 February 2021. Hamilton (ON): McMaster University (developed by Evidence Prime), 2020. Available from gradepro.org.

Green 2018

Green J, Garg S.Annual Research Review: the state of autism intervention science: progress, target psychological and biological mechanisms and future prospects. *Journal of Child Psychology and Psychiatry* 2018;**59**(4):424–43. [DOI: 10.1111/jcpp.12892] [PMID: 29574740]

Greenspan 2007

Greenspan S, Shanker S.The developmental pathways leading to pattern recognition, joint attention, language and cognition. *New Ideas in Psychology* 2007;**25**(2):128-42. [DOI: 10.1016/j.newideapsych.2007.02.007]

Gresham 1990

Gresham FM, Elliott SN.Social Skills Rating System: Manual. Circle Pines (MN): American Guidance Service, 1990.

Gresham 2008

Gresham F, Elliott SN.Social Skills Improvement System (SSIS) Rating Scales. San Antonio (TX): Pearson Education, 2008.

Griffiths 2019

Griffiths S, Allison C, Kenny R, Holt R, Smith P, Baron-Cohen S.The Vulnerability Experiences Quotient (VEQ): a study of vulnerability, mental health and life satisfaction in autistic adults. *Autism Research* 2019;**12**(10):1516-28. [DOI: 10.1002/aur.2162]

Higgins 2002

Higgins JP, Thompson SG.Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539-58. [DOI: 10.1002/sim.1186] [PMID: 12111919]

Higgins 2011

Higgins JP, Altman DG, Sterne JA (editors).Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S (editors). Cochrane Handbook of Systematic Review of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1.

Higgins 2021

Higgins JP, Eldridge S, Li T, editor(s).Chapter 23: Including variants on randomized trials. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

Holck 2004

Holck U.Interaction themes in music therapy: definition and delimitation. *Nordic Journal of Music Therapy* 2004;**13**(1):3-19. [DOI: 10.1080/08098130409478094]

James 2015

James R, Sigafoos J, Green VA, Lancioni GE, O'Reilly MF, Lang R, et al.Music therapy for individuals with autism spectrum disorder: a systematic review. *Review Journal of Autism and Developmental Disorders* 2015;**2**:39-54. [DOI: 10.1007/s40489-014-0035-4]

Kaufman 1987

Kaufman AS, O'Neal MR, Avant AH, Long SW.Introduction to the Kaufman Assessment Battery for Children (K-ABC) for pediatric neuroclinicians. *Journal of Child Neurology* 1987;**2**(1):3-16. [DOI: 10.1177/088307388700200102] [PMID: 3624826]

Kendall 2004

Kendall P, Holmbeck G, Verduin T.Methodology, design, and evaluation in psychotherapy research. In: LambertMJ, editors(s). Bergin and Garfield's Handbook of Psychotherapy and Behavior Change. 5th edition. New York (NY): John Wiley & Sons, 2004:16-43.

Kenny 2016

Kenny L, Hattersley C, Molins B, Buckley C, Povey C, Pellicano E.Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism* 2016;**20**(4):442-62. [DOI: 10.1177/1362361315588200]

Kern 2017

Kern P, Tague DB.Music therapy practice status and trends worldwide: an international survey study. *Journal of Music Therapy* 2017;**54**(3):255-86. [DOI: 10.1093/jmt/thx011]

Kim 1998

Kim H.Gender Differences and Developmental Changes of Adolescents' Self-Esteem [Doctoral dissertation]. Pusan National University, 1998.

Kim 2006

Kim J.The Effects of Improvisational Music Therapy on Joint Attention Behaviours in Children with Autistic Spectrum Disorder [PhD thesis]. Aalborg, DK: Aalborg University, 2006.

Krug 1980

Krug DA, Arick J, Almond P.Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry* 1980;**21**(3):221-9. [DOI: 10.1111/j.1469-7610.1980.tb01797.x] [PMID: 7430288]

Kwak 1998

Kwak K.Self-efficacy and moral justification (1): the developmental pattern. *Korean Journal of Developmental Psychology* 1998;**11**(1):1-11.

Lai 2019

Lai MC, Kassee C, Besney R, Bonato S, Hull L, Mandy W, et al.Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis.. *Lancet Psychiatry* 2019;**6**(10):819-29. [DOI: 10.1016/S2215-0366(19)30289-5] [PMID: 31447415]

Lai 2020

Lai MC, Anagnostou E, Wiznitzer M, Allison C, Baron-Cohen S.Evidence-based support for autistic people across the lifespan: maximizing potential, minimizing barriers, and optimizing the person-environment fit. *Lancet Neurology* 2020;**19**(5):434-451. [DOI: 10.1016/S1474-4422(20)30034-X]

Lambert 2004

Lambert MJ, Ogles BM.The efficacy and effectiveness of psychotherapy. In: LambertMJ, editors(s). Bergin and Garfield's Handbook of Psychotherapy and Behavior Change. 5th edition. New York (NY): John Wiley & Sons, 2004:139-93.

Liberati 2009

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al.The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *British Medical Journal* 2009;**339**:b2700. [DOI: 10.1136/bmj.b2700]

Lipinski 2019

Lipinski S, Blanke ES, Suenkel U, Dziobek I.Outpatient psychotherapy for adults with high-functioning autism spectrum condition: utilization, treatment satisfaction, and preferred modifications. *Journal of Autism and Developmental Disorders* 2019;**49**(3):1154–68. [DOI: 10.1007/s10803-018-3797-1] [PMID: 30415320]

Lord 1994

Lord C, Rutter M, Le Couteur A.Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders* 1994;**24**(5):659-85. [DOI: 10.1007/BF02172145] [PMID: 7814313]

Lord 1999

Lord C, Rutter M, DiLavore PC, Risi S.Autism Diagnostic Observation Schedule: Manual. Los Angeles (CA): Western Psychological Services, 1999.

Lotter 1966

Lotter V.Epidemiology of autistic conditions in young children. *Social Psychiatry* 1966;**1**(3):124-35. [DOI: 10.1007/BF00584048]

Lyall 2017

Lyall K, Croen L, Daniels J, Fallin MD, Ladd-Acosta C, Lee BK, et al.The changing epidemiology of autism spectrum disorders. *Annual Review of Public Health* 2017;**38**:81-102. [DOI: 10.1146/annurev-publhealth-031816-044318] [PMCID: PMC6566093] [PMID: 28068486]

Maenner 2020

Maenner MJ, Shaw KA, Baio J, Washington A, Patrick M, DiRienzo M, et al.Prevalence of autism spectrum disorder among children aged 8 years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *MMWR Surveillance Summaries* 2020;**69**(4):1-12. [DOI: 10.15585/mmwr.ss6904a1] [PMCID: PMC7119644] [PMID: 32214087]

Magyar 2007

Magyar CI, Pandolfi V.Factor structure evaluation of the Childhood Autism Rating Scale. *Journal of Autism and Developmental Disorders* 2007;**37**(9):1787-94. [DOI: 10.1007/s10803-006-0313-9] [PMID: 17437070]

Manly 2016

Manly T, Anderson V, Crawford J, George M, Underbjerg M, Robertson IH.Test of Everyday Attention for Children (TEA-Ch2). 2nd edition. London: Pearson Education, 2016.

Marquez-Garcia 2021

Marquez-Garcia AV, Magnuson J, Morris J, Iarocci G, Doesburg S, Moreno S.Music therapy in autism spectrum disorder: a systematic review. *Review Journal of Autism and Developmental Disorders* 2022;**9**:91-107. [DOI: 10.1007/s40489-021-00246-x]

Marshall 2018

Marshall IJ, Noel-Storr A, Kuiper J, Thomas J, Wallace BC.Machine learning for identifying randomized controlled trials: an evaluation and practitioner's guide. *Research Synthesis Methods* 2018;**9**(4):602-14. [DOI: 10.1002/jrsm.1287] [PMCID: PMC6030513] [PMID: 29314757]

Mayer-Benarous 2021

Mayer-Benarous H, Benarous X, Vonthron F, Cohen D.Music therapy for children with autistic spectrum disorder and/or other neurodevelopmental disorders: a systematic review. *Frontiers in Psychiatry* 2021;**12**:643234. [DOI: 10.3389/fpsyt.2021.643234] [PMCID: PMC8062803] [PMID: 33897497]

McConachie 2015

McConachie H, Parr JR, Glod M, Hanratty J, Livingstone N, Oono IP, et al.Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment* 2015;**19**(41):1-506. [DOI: 10.3310/hta19410] [PMCID: PMC4781156] [PMID: 26065374]

McKenzie 2021

McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV, Thomas J .Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Millstein 2021

Millstein A, Myers-Coffman K, Horowitz S, Kesslick A, Bradt J.Experiences of implementing treatment manuals: clinician, supervisor, and researcher reflections. *Nordic Journal of Music Therapy* 2021;**30**(2):179-91. [DOI: 10.1080/08098131.2020.1790632]

Milton 2012

Milten D.On the ontological status of autism: the 'double empathy problem'. *Disability & Society* 2012;**27**(6):883-7. [DOI: 10.1080/09687599.2012.710008]

Milton 2019

Milton D.Difference versus disability: implications of characterisation of autism for education and support. In: JordanR, RobertsJM, HumeK, editors(s). Handbook of Autism and Education. London (UK): Sage, 2019:3-11.

Moher 2010

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al.CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c869. [DOI: 10.1136/bmj.c869]

Mössler 2019

Mössler K, Gold C, Aßmus J, Schumacher K, Calvet C, Reimer S, et al.The therapeutic relationship as predictor of change in music therapy with young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 2019;**49**(7):2795-809. [DOI: 10.1007/s10803-017-3306-y] [PMID: 28936692]

Mössler 2020

Mössler K, Schmid W, Aßmus J, Fusar-Poli L, Gold C.Attunement in music therapy for young children with autism: revisiting qualities of relationship as mechanisms of change. *Journal of Autism and Developmental Disorders* 2020;**50**(11):3921-34. [DOI: 10.1007/s10803-020-04448-w] [PMCID: PMC7560932] [PMID: 32189227]

Mundy 2003

Mundy P, Delgado C, Block J, Venezia M, Hogan A, Seibert J.A Manual for the Abridged Early Social Communication Scales (ESCS). Coral Gables (FL): University of Miami, 2003.

Muris 1997

Muris P, Steerneman P, Ratering E.Brief report: interrater reliability of the PsychoEducational Profile (PEP). *Journal of Autism and Developmental Disorders* 1997;**27**(5):621-6. [DOI: 10.1023/a:1025838312457] [PMID: 9403376]

Noel-Storr 2020

Noel-Storr AH, Dooley G, Wisniewski S, Glanville J, Thomas J, Cox S, et al.Cochrane Centralised Search Service showed high sensitivity identifying randomised controlled trials: a retrospective analysis. *Journal of Clinical Epidemiology* 2020;**127**:142-50. [DOI: 10.1016/j.jclinepi.2020.08.008] [PMID: 32798713]

Noel-Storr 2021

Noel-Storr AH, Dooley G, Elliott J, Steele E, Shemilt I, Mavergames C, et al.An evaluation of Cochrane Crowd found that crowdsourcing produced accurate results in identifying randomised trials. *Journal of Clinical Epidemiology* 2021;**130**:23-31. [DOI: 10.1016/j.jclinepi.2021.01.006] [PMID: 33476769]

Oldfield 2012

Oldfield A, Bell K, Pool J.Three families and three music therapists: reflections on short term music therapy in child and family psychiatry. *Nordic Journal of Music Therapy* 2012;**21**(3):250-67. [DOI: 10.1080/08098131.2011.640436]

Oono 2013

Oono IP, Honey EJ, McConachie H.Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews* 2013, Issue 4. Art. No: CD009774. [DOI: 10.1002/14651858.CD009774.pub2] [PMID: 23633377]

Ozonoff 2005

Ozonoff S, Goodlin-Jones BL, Solomon M.Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology* 2005;**34**(3):523-40. [DOI: 10.1207/s15374424jccp3403_8] [PMID: 16083393]

Park 2003

Park J, Hoffman L, Marquis J, Turnbull AP, Poston D, Mannan H, et al.Toward assessing family outcomes of service delivery: validation of a family quality of life survey. *Journal of Intellectual Disability Research* 2003;**47**(Pt 4-5):367-84. [DOI: 10.1046/j.1365-2788.2003.00497.x] [PMID: 12787167]

Pasiali 2004

Pasiali V.The use of prescriptive therapeutic songs in a home-based environment to promote social skills acquisition by children with autism: three case studies. *Music Therapy Perspectives* 2004;**22**(1):11-20. [DOI: 10.1093/mtp/22.1.11]

Pasquali 2002

Pasquali L, Wechsler S, Bensusan E.Raven's Colored Progressive Matrices for children: a validation study for Brazil [Matrizes Progressivas do Raven infantil: um estudo de validação para o Brasil]. *Avaliação Psicológica* 2002;**1**(2):95-110. [WEB PAGE: pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712002000200003]

Pater 2017

Pater M, Van Yperen T.Music therapy for children and young people with ASD. An overview of the relevant literature [Muziektherapie voor kinderen en jongeren met ASS. Een overzicht van de relevante literatuur]. *Kind en Adolescent* 2017;**38**:233-59. [DOI: 10.1007/s12453-017-0156-5]

Pereira 2008

Pereira A, Riesgo RS, Wagner MB.Childhood autism: translation and validation of the Childhood Autism Rating Scale for use in Brazil [Autismo infantil: tradução e validação da *Childhood Autism Rating Scale* para uso no Brasil]. *Jornal de Pediatria* 2008;**84**(6):487-94. [DOI: 10.2223/JPED.1828] [PMID: 18923798]

Pickard 2020

Pickard B, Thompson G, Metell M, Roginsky E, Elefant C, Roginsky E."It's not *what's* done, but *why* it's done": Music therapists' understanding of normalisation, maximisation and the neurodiversity movement. *Voices: A World Forum for Music Therapy* 2020;**20**(3):19. [DOI: 10.15845/voices.v20i3.3110]

Pickles 2016

Pickles A, Le Couteur A, Leadbitter K, Salomone E, Cole-Fletcher R, Tobin H, et al.Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *Lancet* 2016;**388**(10059):2501-9. [DOI: 10.1016/ S0140-6736(16)31229-6] [PMCID: PMC5121131] [PMID: 27793431]

Provenzani 2020

Provenzani U, Fusar-Poli L, Brondino N, Damiani S, Vercesi M, Meyer N, et al.What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. *Autism* 2020;**24**(2):274-84. [DOI: 10.1177/1362361319854641] [PMID: 31269800]

Rapin 2008

Rapin I, Goldman S.The Brazilian CARS: a standardized screening tool for autism. *Jornal de Pediatria* 2008;**84**(6):473-5. [DOI: 10.2223/JPED.1843] [PMID: 19060987]

Reschke-Hernández 2011

Reschke-Hernández AE.History of music therapy treatment interventions for children with autism. *Journal of Music Therapy* 2011;**48**(2):169-207. [DOI: 10.1093/jmt/48.2.169] [PMID: 21938891]

RevMan Web 2020 [Computer program]

The Cochrane Collaboration Review Manager Web (RevMan Web).Version 3.1.1. The Cochrane Collaboration, 2020. Available at revman.cochrane.org.

Rimland 1999

Rimland B, Edelson SM.Autism Treatment Evaluation Checklist (ATEC). San Diego (CA): Autism Research Institute, 1999.

Rolvsjord 2005

Rolvsjord R, Gold C, Stige B.Research rigour and therapeutic flexibility: rationale for a therapy manual developed for a randomised controlled trial. *Nordic Journal of Music Therapy* 2005;**14**(1):15-32. [DOI: 10.1080/08098130509478122]

Rosenberg 1965

Rosenberg M.Society and the Adolescent Self-Image. Princeton (NJ): Princeton University Press, 1965.

Rumball 2020

Rumball F, Happé F, Grey N.Experience of trauma and PTSD symptoms in autistic adults: risk of PTSD development following DSM-5 and non-DSM-5 traumatic life events. *Autism Research* 2020;**13**(12):2122-32. [DOI: https://doi.org/10.1002/aur.2306]

Salomon-Gimmon 2019

Salomon-Gimmon M, Elefant C.Development of vocal communication in children with autism spectrum disorder during improvisational music therapy. *Nordic Journal of Music Therapy* 2019;**28**(3):174-92. [DOI: 10.1080/08098131.2018.1529698]

Schopler 1979

Schopler E, Reichler RJ.Individualized Assessment and Treatment for Autistic and Developmentally Disabled Children. Vol. **1. Psychoeducational profile**. Baltimore (MD): Uniersity Park Press, 1979.

Schopler 1980

Schopler E, Reichler RJ, DeVellis RF, Daly K.Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders* 1980;**10**(1):91-103. [DOI: 10.1007/BF02408436] [PMID: 6927682]

Schopler 2010

Schopler E, Van Bourgondien ME, Wellman GJ, Love SR.The Childhood Autism Rating Scale (CARS2). 2nd edition. Los Angeles (CA): Western Psychological Services, 2010.

Schumacher 1994

Schumacher K.Musiktherapie mit Autistischen Kindern. Stuttgart, Jena, New York: G. Fischer , 1994.

Schumacher 1999a

Schumacher K.Musiktherapie und Säuglingsforschung. Frankfurt, DE: Peter Lang, 1999.

Schumacher 1999b

Schumacher K, Calvet-Kruppa C.The "AQR" - an analysis system to evaluate the quality of relationship during music therapy. *Nordic Journal of Music Therapy* 1999;**8**(2):188-91. [DOI: 10.1080/08098139909477974]

Schumacher 2019

Schumacher K, Calvet C, Reimer S.The AQR Tool - Assessment of the Quality of Relationship. Wiesbaden: Reichert Verlag, 2019.

Schünemann 2013

Schünemann H, Brożek J, Guyatt G, Oxman A, editor(s).Handbook for grading the quality of evidence and strength of recommendations using the GRADE approach (updated October 2013). GRADE Working Group, 2013. Available from gdt.guidelinedevelopment.org/app/handbook/handbook.html.

Schünemann 2021

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, et al.Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

Schünemann 2021a

Schünemann HJ, Vist GE, Higgins JP, Santesso N, Deeks JJ, Glasziou P, et al .Chapter 15: Interpreting results and drawing conclusions. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Sharda 2019

Sharda M, Silani G, Specht K, Tillmann J, Nater U, Gold C.Music therapy for children with autism: investigating social behaviour through music. *Lancet. Child & Adolescent Health* 2019;**3**(11):759-61. [DOI: 10.1016/ S2352-4642(19)30265-2] [PMID: 31494080]

Shi 2016

Shi Z-M, Lin G-H, Xie Q.Effects of music therapy on mood, language, behavior, and social skills in children with autism: a meta-analysis. *Chinese Nursing Research* 2016;**3**(3):137-41. [DOI: 10.1016/j.cnre.2016.06.018]

Simpson 2011

Simpson K, Keen D.Music interventions for children with autism: narrative review of the literature. *Journal of Autism and Developmental Disorders* 2011;**41**(11):1507-14. [DOI: 10.1007/s10803-010-1172-y] [PMID: 21203898]

Sinclair 2010

Sinclair J.Cultural commentary: being autistic together. *Disability Studies Quarterly* 2010;**30**(1):no pagination. [DOI: 10.18061/dsq.v30i1.1075]

Sinclair 2012

Ellermann, M.Interview with Jim Sinclair. www.autism.se/RFA/uploads/nedladningsbara%20filer/Interview_with_Jim_Sinclair.pdf (accessed prior to 15 March 2022).

Sparrow 1984

Sparrow SS, Balla DA, Cicchetti DV.Vineland Adaptive Behavior Scales: Interview Edition, Expanded Form. Circle Pines (MN): American Guidance Service, 1984.

Sparrow 1998

Sparrow SS, Balla DA, Cicchetti DV.Vineland Social-Emotional Early Childhood Scales. Manual. Circle Pines (MN): American Guidance Service, 1998.

Stephens 2008

Stephens CE.Spontaneous imitation by children with autism during a repetitive musical play routine. *Autism* 2008;**12**(6):645-71. [DOI: 10.1177/1362361308097117] [PMID: 19005033]

Stern 1985

Stern DN.The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology. New York (NY): Basic Books, 1985. [OCLC: 45430229]

Stern 1989

Stern DN.The representation of relational patterns: developmental considerations. In: SameroffAJ, EmdeRN, editors(s). Relationship Disturbances in Early Childhood: A Developmental Approach. New York (NY): Basic Books, 1989:52-69.

Stern 2010

Stern DN.Forms of Vitality: Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy and Development. Oxford, UK: Oxford University Press, 2010. [DOI: 10.1093/med:psych/9780199586066.001.0001]

Stone 2010

Stone W, Ruble L, Coonrod E, Hepburn S, Pennington M, Burnette C, et al.TRIAD Social Skills Assessment: Assessing Children with Autism Spectrum Disorders. 2nd edition. Nashville (TN): Vanderbilt TRIAD, 2010.

Thaut 2000

Thaut MH.A Scientific Model of Music in Therapy and Medicine. San Antonio, Texas, USA: IMR Press, The University of Texas at San Antonio, 2000.

Thaut 2014

Thaut MH, Gardiner JC.Musical attention control training. In: ThautMH, HoembergV, editors(s). Handbook of Neurologic Music Therapy. Oxford: Oxford University Press, 2014:257-69.

Thomas 2020

Thomas J, McDonald S, Noel-Storr AH, Shemilt I, Elliott J, Mavergames C, et al.Machine learning reduces workload with minimal risk of missing studies: development and evaluation of a randomized controlled trial classifier for Cochrane Reviews. *Journal of Clinical Epidemiology* 2021;**133**:140-51. [DOI: 10.1016/j.jclinepi.2020.11.003] [PMID: 33171275]

Thompson 2012

Thompson G.Family-centered music therapy in the home environment: promoting interpersonal engagement between children with autism spectrum disorder and their parents. *Music Therapy Perspectives* 2012;**30**(2):109-16. [DOI: 10.1093/mtp/30.2.109]

Thorpe 2009

Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al.A PRagmatic-Explanatory Continuum Indicator Summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology* 2009;**62**(5):464-75. [DOI: 10.1016/j.jclinepi.2008.12.011] [PMID: 19348971]

Thye 2018

Thye MD, Bednarz HM, Herringshaw AJ, Sartin EB, Kana RK.The impact of atypical sensory processing on social impairments in autism spectrum disorder. *Developmental Cognitive Neuroscience* 2018;**29**:151-67. [DOI: 10.1016/j.dcn.2017.04.010]

Trevarthen 1999a

Trevarthen C.How music heals. In: WigramT, De BackerJ, editors(s). Clinical Applications of Music Therapy in Developmental Disability, Paediatrics and Neurology. London, UK: Jessica Kingsley Publishers, 1999:7-10.

Trevarthen 1999b

Trevarthen C.Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae* 1999;**3**(1 Suppl):155-215. [DOI: 10.1177/10298649000030S109]

Trevarthen 2011

Trevarthen C.What is it like to be a person who knows nothing? Defining the active intersubjective mind of a newborn human being. *Infant and Child Development* 2011;**20**(1):119-35. [DOI: 10.1002/icd.689]

Varni 1999

Varni JW, Seid M, Rode CA.The PedsQL™: measurement model for the Pediatric Quality of Life Inventory. *Medical Care* 1999;**37**(2):126-39. [DOI: 10.1097/00005650-199902000-00003] [PMID: 10024117]

Warren 2011

Warren Z, Veenstra-VanderWeele J, Stone W, Bruzek JL, Nahmias AS, Foss-Feig JH, et al.Therapies for children with autism spectrum disorders; April 2011. effectivehealthcare.ahrq.gov/sites/default/files/pdf/autism_research.pdf (accessed 18 March 2014).

Wechsler 1949

Wechsler D.Wechsler Intelligence Scale for Children. New York (NY): Psychological Corporation, 1949.

Wechsler 1967

Wechsler D.Manual: Wechsler Preschool and Primary Intelligence Scale. New York (NY): Psychological Corporation, 1967.

Wechsler 1999

Wechsler D.Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio (TX): Psychological Corporation, 1999.

Weissman 1980

Weissman MM, Orvaschel H, Padian N.Children's symptom and social functioning self-report scales. Comparison of mothers' and children's reports. *Journal of Nervous and Mental Disease* 1980;**168**(12):736-40. [DOI: 10.1097/00005053-198012000-00005] [PMID: 7452212]

Weitlauf 2014

Weitlauf AS, McPheeters ML, Peters B, Sathe N, Travis R, Aiello R, et al.Therapies for Children with Autism Spectrum Disorder: Behavioral Interventions Update. Rockville (MD): Agency for Healthcare Research and Quality, 2014. [PMID: 25210724]

Wheeler 2008

Wheeler D, Williams K, Seida J, Ospina M.The Cochrane Library and autism spectrum disorder: an overview of reviews.

Evidence-Based Child Health: A Cochrane Review Journal 2008;**3**(1):3-15. [DOI: 10.1002/ebch.218]

Whipple 2004

Whipple J.Music in intervention for children and adolescents with autism: a meta-analysis. *Journal of Music Therapy* 2004;**41**(2):90-106. [DOI: 10.1093/jmt/41.2.90] [PMID: 15307805]

Whipple 2012

Whipple J.Music therapy as an effective treatment for young children with autism spectrum disorders: a meta-analysis. In: KernP, HumpalM, editors(s). Early Childhood Music Therapy and Autism Spectrum Disorders: Developing Potential in Young Children and their Families. London, UK: Jessica Kingsley Publishers, 2012:58-76.

WHO 1994

World Health Organization.ICD-10: International Statistical Classification of Diseases and Related Health Problems. 10th edition. Geneva (CH): World Health Organization, 1994.

WHO 2021

World Health Organization.ICD-11: International Statistical Classification of Diseases 11th Revision. The Global Standard for Diagnostic Health Information. icd.who.int/en (accessed 5 May 2021).

Wigram 2002

Wigram T.Indications in music therapy: evidence from assessment that can identify the expectations of music therapy as a treatment for autistic spectrum disorder (ASD). Meeting the challenge of evidence based practice. *British Journal of Music Therapy* 2002;**16**(1):11-28. [DOI: 10.1177/135945750201600104]

Wigram 2006

Wigram T, Gold C.Music therapy in the assessment and treatment of autistic spectrum disorder: clinical application and research evidence. *Child: Care, Health and Development* 2006;**32**(5):535-42. [DOI: 10.1111/j.1365-2214.2006.00615.x] [PMID: 16919132]

Wigram 2009

Wigram T, Elefant C.Therapeutic dialogues in music: nurturing musicality of communication in children with autistic spectrum disorder and Rett syndrome. In: MallochS, TrevarthenC, editors(s). Communicative Musicality. New York (NY): Oxford University Press, 2009:423-45.

Winter 2012

Winter P.Loud hands and loud voices. In: BascomJ, editors(s). Loud Hands: Autistic People, Speaking. Autistic Press, 2012:115-28.

References to other published versions of this review

Geretsegger 2014

Geretsegger M, Elefant C, Mössler KA, Gold C.Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews* 2014, Issue 6. Art. No: CD004381. [DOI: 10.1002/14651858.CD004381.pub3] [PMCID: PMC6956617] [PMID: 24936966]

Gold 2003

Gold C, Wigram T.Music therapy for autistic spectrum disorder. *Cochrane Database of Systematic Reviews* 2003, Issue 3. Art. No: CD004381. [DOI: 10.1002/14651858.CD004381]

Gold 2006

Gold C, Wigram T, Elefant C.Music therapy for autistic spectrum disorder. *Cochrane Database of Systematic Reviews* 2006, Issue 2. Art. No: CD004381. [DOI: 10.1002/14651858.CD004381.pub2] [PMID: 16625601]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Arezina 2011

Study characteristics	
Methods	**Allocation:** session order randomised using Latin Square **Blindness:** unclear; random sub-sample (33.33% of sessions) assessed by independent observer **Duration:** 5 weeks **Design:** cross-over
Participants	**Diagnosis:** ASD **N:** 6 **Age range:** 36 to 64 months **Sex:** 5 boys, 1 girl **Setting:** unclear **Location:** Midwestern, USA
Interventions	1.**Experimental (n = 6):** interactive MT; musical instrument play, songs, music books, sung and verbal responses to verbalisations; 6 x 10-minute individual sessions

Arezina 2011 *(Continued)*

	2. **Control (n = 6):** non-music interactive play; non-music toys and books, verbal responses to verbalisations; 6 x 10-minute individual sessions 3. **Control (n = 6):** independent play; 6 x 10-minute individual sessions
Outcomes	Behaviour observation based on videotaped sessions, coded by researcher (33.3% of sessions also coded by another observer) 1. **Social interaction:** requesting or initiating joint attention (number of requests during a given time period), measured during the intervention 2. **Adaptive behaviour:** interaction or engaging in joint attention (percent of 15-second intervals engaged in interaction), measured during the intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Order of sessions (including different therapeutic approaches) was randomised for each child using a Latin Square.
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details about blinding reported; however, a random subsample (33.33%) was assessed by an independent observer (inter-observer agreement ranged from 85.7% to 98.9%).
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Bharathi 2019

Study characteristics

Methods	**Allocation:** allocated by researchers **Blindness:** no blinding **Duration:** 6 months (3 months intervention + 3 months follow-up) **Design:** quasi-experimental, parallel group, with control and a pre-post and follow-up test; single centre, recruited through a convenience sample method

Bharathi 2019 *(Continued)*

Participants	**Diagnosis:** ASD according to DSM-5 criteria; mild, moderate or severe as per Childhood Autism Rating Scales (CARS) scores **N:** 52 **Age range:** 6 to 12 years (mean 9.5) **Sex:** 26 boys, 26 girls **Setting:** unclear **Location:** Coimbatore, South India
Interventions	1.**Experimental (n = 26):** active group MT; singing, dancing, playing musical instruments while listening to music; 3 x 35-minute sessions each week for 3 months 2. **Control (n = 26):** passive group MT; no interaction, only music listening; 3 x 35-minute sessions each week for 3 months
Outcomes	1. **Social interaction:** TRIAD Social Skills Assessment (TSSA) Total, measuring emotion understanding/perspective taking ability, initiating interaction, responding to initiations, maintaining interactions; higher scores are favourable; measured at pre- and post-intervention and at follow-up at 6 months; carried out by researcher considering parent, teacher, and direct interaction with children 2. **Total autism symptom severity:** CARS; lower scores are favourable; measured at pre- and post-intervention; administered by researcher
Notes	**Funding source(s):** Science and Engineering Research Board (SERB) (ECR/2016/001688), Government of India, New Delhi

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"Sixty children (30 boys and 30 girls) from ages 6–12 years were chosen- through a convenience sampling method. [...] The study group was stratified further into two groups as active MT and passive MT intervention group."
Allocation concealment (selection bias)	High risk	As the groups were selected by the researchers, it is likely that the allocation was not concealed.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Non-blinded outcome assessment: "For each child, pre- and post-therapy CARS score was taken by the researcher."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts reported No missing data reported
Selective reporting (reporting bias)	Low risk	No reference to a protocol or trial registration provided. However, as the research received both ethical approval and governmental funding, it is likely that the researchers followed a protocol and reported according to that.
Other bias	Low risk	No financial bias could be found. The researchers might be music therapists. Adequate music therapy method: yes Adequate music therapy training: unclear as no information was provided

Bieleninik 2017

Study characteristics	
Methods	**Allocation:** randomised using computer-generated randomisation list **Blindness:** assessors blind to treatment condition **Duration:** 12 months (5 months intervention + 7 months follow-up) **Design:** parallel-group, multicentre (9 countries, 10 centres)
Participants	**Diagnosis:** ASD **N:** 364 **Age range:** 4.0 to 6.11 years **Sex:** 302 boys, 62 girls **Setting:** outpatient **Location:** Australia, Austria, Brazil, Israel, Italy, Korea, Norway, UK, USA
Interventions	1. **Experimental (n = 90):** individual improvisational music therapy; 30-minute sessions for 5 months, 3 sessions per week 2. **Experimental (n = 92):** individual improvisational music therapy; 30-minute sessions for 5 months, 1 session per week 3. **Control (n = 182):** enhanced standard care
Outcomes	1. **Social interaction:** a. Autism Diagnostic Observation Schedule-Social Affect (ADOS-SA), administered by blinded assessors; lower scores are favourable; evaluated before and during intervention (2 months), post-intervention, and at 7-month follow-up b. Social Responsiveness Scale (SRS) Total, parent-rated (not blinded); lower scores are favourable; evaluated before and during intervention (2 months), post-intervention, and at 7-month follow-up 2. **Quality of life:** 100 mm visual analog scales for parent-reported quality of life of the child (QoL-child) and of the family as a whole (QoL-family); higher scores are favourable; evaluated pre-post intervention, and at 7-month follow-up; completed by participants' parents or guardians 3. **Total autism symptom severity:** ADOS Total, administered by blinded assessors; lower scores are favourable; evaluated before and during intervention (2 months), post-intervention, and at 7-month follow-up 4. **Adverse events:** any adverse events (parent reports) 5. **Adaptive behaviour:** ADOS-Restricted and Repetitive Behaviors (ADOS-RRB), administered by blinded assessors; lower scores are favourable; evaluated before and during intervention (2 months), post-intervention, and at 7-month follow-up
Notes	**Funding source(s):** Research Council of Norway (grant 213844, the Clinical Research and Mental Health Programmes); POLYFON Knowledge Cluster for Music Therapy, The Grieg Academy Department of Music, University of Bergen, Norway; *Australia:* Melbourne Conservatorium of Music, the University of Melbourne; *Austria:* Danish Council for Independent Research/Humanities (FKK), Aalborg University, and Faculty of Psychology, University of Vienna; *Brazil:* Research Incentive Fund (FIPE) of the Hospital de Clínicas de Porto Alegre (HCPA); *Italy:* IRCCS Stella Maris Foundation, Pisa, and University of Pisa; *Korea:* Research Institute of Health and Science, Jeonju University; *Norway:* Health Authority of Western Norway (Helse Vest grant 911800), Bergen municipality (Bergen Culture School), Fjell municipality (Fjell Culture School); *UK:* National Institute for Health Research (Health Technology Assessment Programme, National Institute for Health Research grant 12/167/95), Cambridge and Peterborough Foundation National Health

Bieleninik 2017 *(Continued)*

	Service Trust, The Evelyn Trust, Cambridgeshire Music, Anglia Ruskin University; *USA:* Molloy College Faculty research scholarship and a collaborative research grant from the Mid-Atlantic Region of the American Music Therapy Association

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated sequence
Allocation concealment (selection bias)	Low risk	Randomisation list stored centrally, allocation via an electronic system after decision on inclusion
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessors of the primary outcome were blinded and success of blinding was verified. SRS was administered by parents which were not blinded to intervention.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Low attrition rate
Selective reporting (reporting bias)	Low risk	Trial was prospectively registered; trial protocol published; original trial protocol included with trial report
Other bias	Low risk	No personal or financial bias could be found. The majority of researchers involved are trained music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

Brownell 2002

Study characteristics

Methods	**Allocation:** quasi-randomised, possibly randomised ('counterbalanced') **Blindness:** independent assessor (teacher), blinding not reported **Duration:** 4 weeks **Design:** cross-over
Participants	**Diagnosis:** autism **N:** 4 **Age range:** 6 to 9 years **Sex:** 4 boys, no girls **Setting:** elementary school **Location:** Eastern Iowa, USA
Interventions	1. **Experimental (n = 4):** structured receptive MT; songs with social stories; 5 individual, daily sessions 2. **Experimental (n = 4):** structured receptive 'story therapy'; reading of social stories; 5 individual, daily sessions

Brownell 2002 *(Continued)*

	3. **Control (n = 4):** no intervention; 2 x 5 days
Outcomes	1. **Adaptive behaviour:** frequency of repetitive behaviours outside therapy sessions (in classroom); occurrence of behaviour was assessed by independent observers (i.e. teachers; inter-rater reliability 0.86-0.94) during intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Assignment to a counterbalanced treatment order (either ABAC or ACAB); unclear whether participants were randomly assigned to the two different treatment orders
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Outcomes were assessed by a teacher or instructional associate assigned to the participant. No details given on blinding of assessors
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcomes (targeted behaviours) of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Buday 1995

Study characteristics

Methods	**Allocation:** randomised **Blindness:** assessor blind to the nature of the hypothesis and to treatment condition **Duration:** 2 weeks **Design:** cross-over
Participants	**Diagnosis:** autism **N:** 10 **Age range:** 4 to 9 years **Sex:** 8 boys, 2 girls **Setting:** public school **Location:** Chicago, Illinois, USA

Buday 1995 *(Continued)*

Interventions	1. **Experimental (n = 10):** structured receptive MT; songs used to teach signs; 5 individual sessions 2. **Control (n =10):** 'Rhythm therapy'; rhythmic speech used to teach signs; 5 individual sessions
Outcomes	Imitating behaviour in sessions (rating of a video recording with sound turned off to ensure blinding of rater; rater blind to nature of hypothesis; inter-rater agreement 98% for 25% of each participant's scores) 1. **Non-verbal communication:** sign imitation, assessed during intervention 2. **Verbal communication:** speech imitation, assessed during intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Counterbalancing of target signs for each treatment condition. "Five of the subjects were randomly assigned to be tested on the music condition during the first week, while the other five were tested on the rhythm condition."
Allocation concealment (selection bias)	Unclear risk	No details given, probably no
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessments were conducted by a person blinded to the nature of the hypothesis. Unlikely that the researcher was blinded to the treatment condition as assessments were based on video analysis of the treatment sessions.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Unclear risk	No personal or financial bias could be found. Adequate music therapy method: unclear Adequate music therapy training: unclear

Chen 2010

Study characteristics

Methods	**Allocation:** randomly divided into experimental group and control group (no further specification) **Blindness:** not reported **Duration:** 3 months **Design:** parallel group
Participants	**Diagnosis:** childhood autism according to DSM-IV **N:** 30

Chen 2010 *(Continued)*

	Age range: 2 to 6 years **Sex:** 27 boys, 3 girls **Setting:** Ninth People's Children's Hospital **Location:** Chongqing, China
Interventions	1. **Experimental (n = 15):** comprehensive treatment (medicine and education, including auditory integration training, sensory integration training, special education, language therapy, play therapy, etc.) plus MT; 4 times a week for 30 minutes each, for 3 months, including receptive and active methods 2. **Control (n = 15):** comprehensive treatment alone; 3 months. The process, specifically: the therapist plays the opening song, the children listen, rhythm, experience the rhythm, perform; sing for a total of 5 minutes; the therapist will improvise to guide the children to sing, recreate and knock; play for 10 minutes; rest for 5 minutes; the therapist plays musical stories and guides the children to act, sing and make corresponding expressions and actions for 5 mins; the therapist plays the goodbye song, children listen, move, experience rhythm, and children sing for a total of 5 minutes.
Outcomes	1. **Total autism symptom severity:** Autism Child Behavior Scale (AuBC) Total; lower scores are favourable; completed by parents pre- and post-intervention 2. **Adaptive behaviour:** Psychoeducational Profile (PEP); higher scores are favourable; rated by researchers pre- and post-intervention
Notes	**Funding source(s):** Chongqing Natural Science Foundation (CSTC, 2009BB5129)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomly divided into experimental group and control group (no further specification)
Allocation concealment (selection bias)	Unclear risk	No details provided
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details given on blinding of assessors
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported.
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol. It was likely that all intended measures were included in the analysis.
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: yes (音乐治疗由专门的音乐治疗师进行训练" translates as "music therapy was provided by a professional music therapist")

Chen 2013

Study characteristics	
Methods	**Allocation:** randomised (no further specification) **Blindness:** not reported **Duration:** 3 months **Design:** parallel group
Participants	**Diagnosis:** childhood autism based on DSM-IV criteria **N:** 27 **Age range:** 4 to 5 years **Sex:** 27 boys, no girls **Setting:** Ninth People's Hospital **Location:** Chongqing, China
Interventions	1. **Experimental (n = 9):** social stories MT (including learning to sing social story songs, performing social story content, etc.), in addition to standard care; sessions twice a week, for 50 minutes each, over 3 months 2. **Experimental (n = 9):** MT without lyrics (music group, including learning the tune of the song learned by the social story group, the lyrics replaced by meaningless pronunciation such as "ah.."), in addition to standard care; sessions twice a week, for 50 minutes each, over 3 months 3. **Control (n = 9):** standard care; "medicine and education", including auditory integration training, sensory integration training, special education training, speech therapy, play therapy, etc.
Outcomes	1. **Social interaction:** Childhood Autism Rating Scale (CARS) social communication domain; before treatment and after 3 months (end of treatment); lower scores are favourable; assessed by specially trained clinicians pre- and post-intervention 2. **Non-verbal communication:** CARS non-verbal communication, assessed by clinicians pre- and post-intervention 3. **Verbal communication:** CARS verbal communication, assessed by clinicians pre- and post-intervention 4. **Total autism symptom severity:** Autism Child Behavior Scale (AuBC) Total; lower scores are favourable; before treatment and after 3 months (end of treatment); assessed by caregiver pre- and post-intervention
Notes	**Funding source(s):** Chongqing Natural Science Foundation (CSTC, 2009BB5129); Chongqing Medical Specialty Construction (Chongqing Health Science and Education 2009, 71)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised (without further specification)
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	The scale was assessed by specially trained clinicians. No details given on blinding of assessors

Chen 2013 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol. It was likely that all intended measures were included in the analysis.
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: unclear, however it is likely that it was also a professional music therapist, as described in Chen 2010, who provided the music therapy.

Farmer 2003

Study characteristics	
Methods	**Allocation:** randomised **Blindness:** not known **Duration:** 5 days **Design:** parallel group
Participants	**Diagnosis:** autism **N:** 10 **Age range:** 2 to 5 years **Sex:** 9 boys, 1 girl **Setting:** homes and therapy centres **Location:** Atlanta, Georgia, USA
Interventions	1. **Experimental (n = 5):** MT sessions; combined active and receptive: guitar playing, songs, mostly individual sessions of 20 minutes 2. **Control (n = 5):** placebo; no music sessions
Outcomes	Responses within sessions (inter-rater agreement 91% for independent observer who rated 10% of sessions) 1. **Non-verbal communication:** gestural responses, assessed during intervention 2. **Verbal communication:** verbal responses, assessed during intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomised, no further details given
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias)	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias.

Farmer 2003 *(Continued)*

All outcomes		The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Unlikely that assessors were masked to the randomisation result as the assessments were based on video analysis of the sessions
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Gattino 2011

Study characteristics	
Methods	**Allocation:** balanced randomisation using a table of random numbers **Blindness:** assessors blind **Duration:** 7 months **Design:** parallel group
Participants	**Diagnosis:** ASD **N:** 24 **Age range:** 7 to 12 years (mean 9.75 years) **Sex:** 24 boys, no girls **Setting:** hospital **Location:** Porto Alegre, Brazil
Interventions	1. **Experimental (n = 12)**: relational MT; improvisation not using a structured protocol; 3 assessment sessions, 16 intervention sessions, 1 final assessment session, in addition to standard treatment, 20 x 30-minute individual sessions, scheduled weekly 2. **Control (n = 12)**: standard treatment; clinical routine activities, including medical examinations and consultations
Outcomes	1. **Social interaction:** Childhood Autism Rating Scale, Brazilian version (CARS-BR), social communication domain; lower scores are favourable; carried out by assessor blind to group allocation pre- and post-intervention 2. **Non-verbal communication:** CARS-BR non-verbal communication domain; lower scores are favourable; carried out by assessor blind to group allocation pre- and post-intervention 3. **Verbal communication:** CARS-BR verbal communication domain; lower scores are favourable; carried out by assessor blind to group allocation pre- and post-intervention
Notes	**Funding source(s):** Fund of Incentive to Research of Porto Alegre Clinical Hospital (project no. 08006), Brazilian Research Council (CNPq)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Gattino 2011 *(Continued)*

Random sequence generation (selection bias)	Low risk	Randomised (computer-generated random sequence)
Allocation concealment (selection bias)	Low risk	Allocation was conducted by an external investigator.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessors were blinded to the randomisation result.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Ghasemtabar 2015

Study characteristics	
Methods	**Allocation:** not randomised (allocated by researcher, matched by age and gender) **Blindness:** no blinding **Duration:** 3.5 months (45 days intervention + 2 months follow-up) **Design:** parallel group, pretest/post-test/follow-up
Participants	**Diagnosis:** mild to moderate autism based on Childhood Autism Rating Scale (CARS) scores **N:** 27 **Age range:** 7 to 12 years (mean 9.1) **Sex:** 14 boys, 13 girls **Setting:** three child and adolescent psychiatry centres **Location:** Teheran, Iran
Interventions	1. **Experimental (n = 13):** MT Orff-Schulwerk; 12 sessions (2 sessions of 1 h/week) in 45 days, delivered by two music therapists, including music listening, singing songs and chants, clapping, movement and dancing, special musical drama, playing of instruments, etc. 2. **Control (n = 14):** without intervention
Outcomes	1. **Social interaction:** Social Skills Rating System Scale-Parent form (SSRS-P); high score means high social skills; parent-rated, pre- and post-intervention and 2 months after the intervention
Notes	**Funding source(s):** none
Risk of bias	

Ghasemtabar 2015 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"Regarding the fact that the present research's design is pretest/post-test-follow–up with control group, in order to eliminate the possible intervening variables, we have tried to match the children of both groups by age and gender variables. Therefore, 6 girls and 7 boys were replaced in the experiment group (n = 13), and 7 girls and 7 boys were replaced in the control group (n = 14)."
Allocation concealment (selection bias)	High risk	No randomisation procedures followed
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	The SSRS is a parent-administered outcome. It was unlikely that parents were blinded to intervention.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	No published study protocol/trial registration. It was likely that all intended measures were included in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Huang 2015

Study characteristics	
Methods	**Allocation:** randomised (no further specification) **Blindness:** not reported **Duration:** 2 months **Design:** parallel group
Participants	**Diagnosis:** autism **N:** 60 **Age range:** 2 to 9 years **Sex:** 47 boys, 13 girls **Setting:** Xihua County People's Hospital **Location:** Zhoukou, China
Interventions	1. **Control (n = 30):** integrated therapy; mainly including auditory integration training, speech therapy, sensory integration training, play therapy, special education training, etc. 2. **Experimental (n = 30):** integrated therapy plus MT; as described above and including listening, active participation, improvisation, suggestive relaxation, etc.

Huang 2015 *(Continued)*

	6 sessions per week for 3 weeks (18 sessions), 30-35 minutes each, then 10 days rest, followed by another period of 3 weeks with 6 sessions each (18 more sessions), followed by another period of rest until day 60
Outcomes	1. **Total autism symptom severity:** Autism Child Behavior Scale (AuBC) Total, including emotional, social, behavioural, and perceptual aspects; low scores are favourable; rated by parents/caregivers pre- and post-intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised (without further specification)
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details given on blinding of assessors
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	No published study protocol/trial registration. It was likely that all intended measures were included in the analysis.
Other bias	Unclear risk	No financial or personal bias could be found. Adequate music therapy method: yes Adequate music therapy training: not clear whether the intervention was conducted by a music therapist. It only reported "therapist" (治疗师) in the paper.

Kim 2008

Study characteristics

Methods	**Allocation:** randomised **Blindness:** assessors blind to the treatment condition, except for parent-based measures conducted by mothers **Duration:** 8 months (approximately 4 months per intervention) **Design:** cross-over
Participants	**Diagnosis:** autism **N:** 15 at entry; 10 for analysis

Kim 2008 *(Continued)*

	Age range: 39 to 71 months (mean 51 months) **Sex:** 13 boys, 2 girls at entry; 10 boys; no girls for analysis **Setting:** private practice clinic **Location:** Seoul, Korea
Interventions	1. **Experimental (n = 10):** improvisational MT; 12 x 30-minute individual sessions, scheduled weekly 2. **Control (n = 10):** play sessions with toys; 12 x 30-minute sessions, scheduled weekly
Outcomes	1. **Social interaction:** a. Social Approach Subscale (Pervasive Developmental Disorder Behavior Inventory, PDDBI); high scores are favourable; completed by parents (not blind) and independent observers (blinded) before, during, and immediately after the intervention b. behavioural observations in sessions (DVD recordings), assessed by researchers during intervention (inter-rater reliability, with rater blind to session order for 30% of recordings, between 0.59 and 0.98): i. turn-taking frequency and duration ii. initiation of engagement frequency iii. imitation frequency iv. emotional synchronicity frequency and duration v. musical synchronicity frequency and duration vi. joy frequency and duration 2. **Non-verbal communication:** a. Early Social Communication Scale (ESCS), abridged version; high scores are favourable; administered by assessors blind to group allocation pre- and post-intervention b. behavioural observations in sessions: eye contact frequency and duration 3. **Adaptive behaviour:** behavioural observations in sessions: a. compliant response frequency b. non-compliant response frequency c. no response frequency 4. **Quality of family relationships:** Mother Play Intervention Profile (MPIP); video observations of play situations at home scored by researcher, combined with independent observer's scores for a third of the sessions; assessed pre- and post-intervention
Notes	**Funding source(s):** Aalborg University, Denmark

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised (picking the randomisation result from an opaque box)
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Assessors were blinded to the randomisation result, except for non-generalised measures, ESCS, and MPIP, where a random subsample (30%) was additionally assessed by independent observers (inter-rater reliability ranging from 0.70 to 0.98)

Kim 2008 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	High dropout rate (5 of 15 participants dropped out) Data from dropouts were excluded.
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researchers are a music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

LaGasse 2014

Study characteristics	
Methods	**Allocation:** randomised using computerised random numbers table; placed in a group in the order that they consented to the study **Blindness:** no blinding **Duration:** 8 weeks (5 weeks intervention + 3 weeks follow-up) **Design:** parallel group, single centre
Participants	**Diagnosis:** ASD; negative report of dual disability diagnosis **N:** 17 **Age range:** 6 to 9 years (mean 7.58) **Sex:** 13 boys, 4 girls **Setting:** Colorado State University **Location:** Fort Collins, USA
Interventions	1. **Experimental (n = 9):** MT group; twice a week for 50 minutes, over 5 weeks; small groups (3-4 children/group) led by board-certified music therapist (the Transformational Design Model (Thaut 2000) was used to create music experiences that were functionally similar to the non-musical experiences, with the addition of music experiences and cues to facilitate the desired social skills. The primary role of the music was to provide anticipatory cues to aid in follow-through with all tasks and to use engagement in music-making to practice the social skills. Rhythmic cues and music structure were used to help the child plan their response, anticipate the timing for the response, and follow through with the response). 2. **Control (n = 8):** social skills group; twice a week for 50 minutes, over 5 weeks; included cooperative play experiences that involved taking turns, passing cards/game pieces, and interacting with their peers. During these activities, the lead therapist provided cues and prompts to facilitate peer-to-peer interaction and joint attention to manipulatives/peers.
Outcomes	1. **Social interaction:** Social Responsiveness Scale (SRS) Total; lower scores are favourable; completed by child's parent/caregiver before and after intervention 2. **Non-verbal communication:** behavioural observations at 3rd and 10th session, coded by trained music therapy research assistants: a. eye gaze b. joint attention with child c. withdrawal behaviours d. joint attention with adult e. initiation of communication with child f. initiation of communication with adult g. response to communication 3. **Total autism symptom severity:** Autism Treatment Evaluation Checklist (ATEC) Total, investigated four areas: speech and communication, sociability, sensory/cognitive awareness, health/physical be-

LaGasse 2014 *(Continued)*

	haviour; lower scores indicate higher functioning (i.e. lower scores are favourable); completed by parents (before study, after sessions 2, 4, and 6, and 3 days following the final session, and at follow-up 3 weeks after session completion), and by the lead therapist (after sessions 2, 4, 8, and 10)
Notes	**Funding source(s):** Arthur Flagler Fultz Research Fund of the American Music Therapy Association

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised using computerised random numbers table; placed in a group in the order that they consented to the study
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Non-blinded. The SRS and the ATEC are parent-administered. Parents were not blinded to the treatment allocation. Assessors were blinded to the order of sessions for the rating of behaviour via video observation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Not all the data from randomised participants were included in the published report. However, in our analysis they were included.
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol, however ethical approval was provided. It was likely that the researcher followed the procedures that were ethically approved.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Lim 2010

Study characteristics

Methods	**Allocation:** randomised **Blindness:** assessors blind to the purpose of the study **Duration:** 5 days **Design:** parallel group
Participants	**Diagnosis:** ASD **N:** 50 **Age range:** 3 to 5 years (mean 4.8 years) **Sex:** 44 boys, 6 girls **Setting:** recruiting site (schools, therapy centres, etc.) **Location:** USA

Lim 2010 *(Continued)*

Interventions	1. **Experimental (n = 18):** music training; 'Developmental Speech and Language Training through Music'; videotaped songs with target words; 6 individual sessions within 3 days 2. **Control (n = 18):** speech training; videotaped spoken stories with target words; 6 individual sessions within 3 days 3. **Control (n = 14):** no training
Outcomes	1. **Verbal communication:** Verbal Production Evaluation Scale (VPES); behaviour observation based on videotaped post-test sessions, coded by two trained speech/language pathologists specialised in treating young children with language impairment who were blind to the purpose of the study (inter-rater reliability 0.999)
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised, no further details given
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessors were blind to the purpose of the study.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Lim 2011

Study characteristics

Methods	**Allocation:** training order randomised **Blindness:** assessors blind to the purpose of the study **Duration:** 2 weeks **Design:** cross-over
Participants	**Diagnosis:** ASD

Lim 2011 *(Continued)*

	N: 22 **Age range:** 3 to 5 years (mean 4.3 years) **Sex:** 17 males, 5 females **Setting:** no details given **Location:** USA
Interventions	1. **Experimental (n = 22):** music training ("music incorporated Applied Behavior Analysis Verbal Behavior"; sung instructions, songs with target words); 6 individual sessions within 2 weeks 2. **Control (n = 22):** speech training; applied behaviour analysis verbal behaviour; spoken instructions, sentences with target words; 6 individual sessions within 2 weeks 3. **Control (n = 22):** no training
Outcomes	1. **Verbal communication:** Verbal Production Evaluation Scale (VPES); behaviour observation of videotaped post-test sessions; coded by two trained speech/language pathologists specialised in treating young children with language impairment who were blind to the purpose of the study
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Order of sessions (including different therapeutic approaches) was randomised for each child using a random number chart. Participants were also randomly assigned which order to receive the training to avoid order effects.
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Assessors were blind to the purpose of the study.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researchers are music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

Mateos-Moreno 2013

Study characteristics

Methods	**Allocation:** group allocation conducted incidentally by care-centre staff

Mateos-Moreno 2013 *(Continued)*

	Blindness: not reported **Duration:** 17 weeks **Design:** parallel group
Participants	**Diagnosis:** severe autism (DSM-IV); Childhood Autism Rating Scale (CARS) ≥ 37 **N:** 16 **Age range:** not reported (mean 25 years) **Sex:** 15 males, 1 female **Setting:** specialised care centre **Location:** Spain
Interventions	1. **Experimental (n = 8):** combined dance/movement and MT; total of 36 ~ 60-minute group sessions (2 per week), led by both a music and dance accredited therapist. Plurisensory approach, with musical activities related to the Orff method and instruments: patients sitting in a circle beat a tempo with percussive instruments imitatively and creatively with or without background music; song tunes in a limited tessitura; corporal percussion and dancing while singing; and 'gesturalised' song and lyric meanings/feelings. Background classical music always present while patients were entering, sitting and leaving the room. Activities: massage the classmate with a small ball; imagine and simulate situations; imitate or guess emotions showed in pictures; move on the ground in different positions; role-play; dancing; drawing; moving as a 'flamenco' dancer; playing with objects 2. **Control (n = 8)**: TAU; no alternative therapies apart from regular activities
Outcomes	1. **Total autism symptom severity:** Revised Clinical Scale for the Evaluation of Autistic Behavior (ECA-R) global score; lower scores are favourable; rated by two independent experienced clinician psychologists every three weeks during intervention and immediately after the intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Group allocation conducted incidentally by care centre staff
Allocation concealment (selection bias)	High risk	Not concealed
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that individuals with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details provided; "participants were monitored by two independent psychologists."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No personal and financial bias could be found. Adequate music therapy method: yes

Mateos-Moreno 2013 *(Continued)*

	Adequate music therapy training: yes

Moon 2010

Study characteristics	
Methods	**Allocation:** unclear **Blindness:** no blinding (self-report by participants) **Duration:** 9 weeks **Design:** parallel group, single centre
Participants	**Diagnosis:** Asperger's syndrome based on DSM-IV criteria **N:** 20 **Age range:** 9 to 10 years **Sex:** 14 boys, 6 girls **Setting:** hospital **Location:** Seoul, Korea
Interventions	1. **Experimental (n = 10):** music drama; based on Theory of Mind approach, including narration, song, and musical instrument play (the content of a story/fairy tale and the actions and emotions of the characters are enacted using music as the main medium); 2 sessions per week, 18 sessions in total, 40 minutes each 2. **Control (n = 10):** story-sharing/story-telling activity through talking/listening and discussing the contents; story/fairy tale is told/enacted in an interactive way, using nonverbal elements such as emotion-related facial expressions and gestures; 2 sessions per week, 18 sessions in total, 40 minutes each
Outcomes	1. **Identity formation** a. Self-Awareness Scale of Fenigstein (Fenigstein 1979), standardised in Korean by Hyukki Eun (Eun 2000); high scores mean high degree of self-awareness; self-report by participants pre- and post-intervention b. Self-Efficacy Scale by Bandura (Bandura 1978), applied to Korean by Keumjoo Kwak (Kwak 1998); high score means high degree of self-efficacy; self-report by participants pre- and post-intervention c. Self-Esteem Scale by Rosenberg (Rosenberg 1965), revised and supplemented by Heewha Kim (Kim 1998) and re-constructed by Youngsook Cho (Cho 2004); high score means high self-esteem; self-report by participants pre- and post-intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	No information given
Allocation concealment (selection bias)	High risk	No information given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias)	High risk	The chosen outcomes on self-esteem/awareness/efficacy were administered by the participants.

Moon 2010 *(Continued)*
All outcomes

Incomplete outcome data (attrition bias) All outcomes	Low risk	Likely to be no dropouts or missing data
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Unclear risk	No personal and financial bias could be found. Adequate music therapy method: unclear (music drama) Adequate music therapy training: unclear

Porter 2017

Study characteristics	
Methods	**Allocation:** randomised **Blindness:** single blind **Duration:** 6 months (12 weeks intervention + follow-up at 26 weeks after randomisation) **Design:** parallel group, multicentre
Participants	**Diagnosis:** ASD based on ICD-10 criteria **N:** 47 **Age range:** 8 to 16 years **Sex:** 34 boys, 13 girls **Setting:** 6 community care Child and Adolescent Mental Health Services (CAMHS) within the Belfast Health and Social Care Trust **Location:** Belfast, Northern Ireland
Interventions	1. **Experimental (n = 24):** MT; Alvin model of 'Free Improvisation' plus TAU; 12 weekly individual sessions of 30 minutes 2. **Control (n = 23):** TAU; psychiatric counselling or medication, or both
Outcomes	1. **Social interaction:** Social Skills Improvement System Rating Scales (SSIS) measuring communicative and interactional skills; high scores are favourable; parental report and self-report pre-post intervention and 13 weeks after the intervention 2. **Quality of life:** EQ5D Health State Scale; completed by a parent pre- and post-intervention; not sufficiently completed for analyses due to participant fatigue 3. **Adverse events:** All serious adverse events (SAEs) were to be recorded in the case report form during the study and reported to the Principal Investigator within 24 hours. They were then to be reviewed by the Trial Steering Committee (TSC) at regular intervals throughout the trial. The Principal Investigator together with the trial sponsor collated and reported annual safety reports to the Trial Steering Committee. The trial sponsor and the chair of the ethics committee were also informed about severe adverse events by the TSC chair if considered appropriate. 4. **Adaptive behaviour:** Child Behavior Checklist (CBC) measuring social functioning; low scores are favourable; completed by a parent pre- and post-intervention and 13 weeks after the intervention 5. **Quality of family relationships:** Family Assessment Device (FAD) measuring family functioning; low scores are favourable; completed by a parent pre- and post-intervention and 13 weeks after the intervention 6. **Identity formation:** Rosenberg Self-Esteem Scale; high scores are favourable; completed by the young person pre- and post-intervention and at 13-week follow-up 7. **Depression:** Centre for Epidemiological Studies Depression Scale for Children (CES-DC); low scores are favourable; completed by the young person pre- and post-intervention and at 13-week follow-up

Porter 2017 *(Continued)*

Notes	**Funding source(s):** Big Lottery Fund

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The randomisation list was computer-generated, using nQuery Advisor v. 16.01, stratified by centre with a 1:1 allocation using random variable block lengths of 2, 4, 6 and 8.
Allocation concealment (selection bias)	Low risk	Randomisation was carried out independently.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that individuals with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Non-blinded. Outcomes were measured via parental- and self-reports.
Incomplete outcome data (attrition bias) All outcomes	Low risk	"In line with the intention-to-treat principle, patients who attended fewer sessions were not excluded from data analysis."
Selective reporting (reporting bias)	Low risk	Published study protocol. All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: yes

Rabeyron 2020

Study characteristics

Methods	**Allocation:** randomised (using a generated randomisation list) **Blindness:** assessors blind to treatment condition **Duration:** 8 months **Design:** parallel group
Participants	**Diagnosis:** ASD based on Childhood Autism Rating Scale (CARS) **N:** 37 **Age range:** 4 to 7 years **Sex:** 32 boys, 5 girls **Setting:** 5 psychiatric day-care centres **Location:** Nantes, France
Interventions	1. **Experimental (n = 19):** MT; 25 weekly, structured, 30-minute sessions (5 minutes opening ritual of music listening, 20 minutes instrumental and vocal improvisation, 5 minutes closing ritual of music listening) performed by a music therapist and a co-therapist (nurse or educator)

Rabeyron 2020 *(Continued)*

	2. **Control (n = 18):** music listening; 25 weekly, 30-minute sessions performed by a nurse or educator, listening to playlist of commercial music, including French and foreign songs
Outcomes	1. **Global improvement:** Clinical Global Impression (CGI); higher scores indicating greater severity; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention 2. **Social interaction:** CARS, social communication domain; lower scores are favourable; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention 3. **Non-verbal communication:** CARS non-verbal communication domain; lower scores are favourable; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention 4. **Verbal communication:** CARS verbal communication domain; lower scores are favourable; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention 5. **Total autism symptom severity:** CARS Total; lower scores are favourable; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention 6. **Adaptive behaviour:** Aberrant Behavior Checklist (ABC); low scores are favourable; assessed by two trained clinical psychologists who were blind to group allocation pre- and post-intervention
Notes	**Funding source(s):** Entreprendre pour Aider; Academie Francaise

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Children were randomly assigned to one of these two groups using a generated randomization list for each group at t0."
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"[A]ll participants were assessed by two trained clinical psychologists working at the Nantes Hospital. They were totally blind, both at t0 and t1, to the groups the children belonged to".
Incomplete outcome data (attrition bias) All outcomes	Low risk	All available data were included.
Selective reporting (reporting bias)	Low risk	Outcome measures matched the evaluation aims in the trial registration.
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: yes

Sa 2020

Study characteristics

Sa 2020 *(Continued)*

Methods	**Allocation:** randomised - no further information provided; participants stratified by severity level (mild, moderate, severe) **Blindness:** not blinded (outcome measures and treatment protocol applied by the same person) **Duration:** 7-11 weeks **Design:** parallel group, single centre
Participants	**Diagnosis:** ASD; severity check at baseline administered by the PI using the CARS2-HF **N:** 23 **Age range:** 10 to 14 years (mean 12.13 years) **Sex:** 16 males, 7 females **Setting:** school **Location:** Central Valley, California, USA
Interventions	1. **Experimental (n = 11):** Music attention control training (MACT); 45-minute group sessions including 5 to 6 participants led by a board-certified music therapist. MACT (Thaut 2014) includes structured active or receptive musical exercises involving precomposed performance or improvisation in which musical elements (pitch, rhythm, dynamics, etc.) cue different musical responses to address attention skills (selective, sustained, and alternating functions). 2. **Control (n = 12):** waiting-list control group
Outcomes	1. **Cognitive ability:** Test of Everyday Attention for Children 2 (TEA-Ch 2): selective attention, sustained attention, switching attention; administered by the therapist pre- and post-intervention, therefore not used in meta-analysis
Notes	**Funding sources(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that individuals with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Non-blinded. Outcome assessment was performed by the therapist.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.

Sa 2020 *(Continued)*

Other bias	Low risk	No personal and financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: yes

Schwartzberg 2013

Study characteristics

Methods	**Allocation:** randomised by the principal investigator (PI) **Blindness:** no blinding **Duration:** 3 days **Design:** parallel, 6-group, randomised design with cluster-randomisation
Participants	**Diagnosis:** ASD **N:** 30 for analysis **Age range:** 9 to 21 years (mean 15.79) **Sex:** 29 boys, 1 girl **Setting:** 3 separate 1-week summer camps for ASD **Location:** Minneapolis, Minnesota, USA
Interventions	1. **Experimental (n = 16)**: MT structured plus music-based social story; 50-minute sessions for 3 consecutive days. Consisted of 6 campers and 8 staff. MT sessions followed a similar routine (hello song, movement and music intervention, social story intervention, instrument playing intervention, relaxation and music, and goodbye song) 2. **Control (n = 14)**: social story; sessions for 3 consecutive days
Outcomes	1. **Social interaction:** Autism Social Skills Profile (ASSP; Bellini 2007); high scores are favourable; completed by parents post-intervention (one week after the end of the camp) 2. **Verbal communication:** comprehension checks; administered by camp counsellors, completed by participants before and one week after intervention and at the end of each day of intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"The PI cluster-randomised participants". No further details about randomisation method given
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that individuals with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Non-blinded. Outcomes were measured via parental- and self-reports.

Schwartzberg 2013 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	Only data from those also completing the follow-up assessment were included.
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol; however ethical approval. It was likely that all outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researchers are music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

Schwartzberg 2016

Study characteristics	
Methods	**Allocation:** cluster-randomisation performed by the principal investigator (PI) **Blindness:** no blinding **Duration:** 3 days **Design:** parallel group, cluster-randomised
Participants	**Diagnosis:** ASD **N:** 29 for analysis **Age range:** 9 to 21 years (mean 15.57) **Sex:** 26 boys, 3 girls **Setting:** 3 separate 1-week summer camps **Location:** Minnesota, USA
Interventions	1. **Experimental (n = 13):** MT; sing short stories (music-based social story session). Procedure: During each day of the summer camp, the PI provided 50-minute music sessions to all campers attending the camp. Each music session consisted of six campers and eight staff. 2. **Control (n = 16):** read aloud short stories
Outcomes	**Verbal communication:** comprehension checks; administered by camp counsellors, completed by participants each day of intervention at the conclusion of the MT session (short-term) and approximately seven hours later (long-term)
Notes	**Funding source(s):** not reported

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	"Participants were cluster-randomised". No further details about randomisation method given
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that individuals with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.

Schwartzberg 2016 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	High risk	Non-blinded. Outcomes self-administered by participants
Incomplete outcome data (attrition bias) All outcomes	High risk	Only data from those also completing the follow-up assessment were included.
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol. It was likely that there were no deviations from the intended intervention that arose because of the trial context.
Other bias	Low risk	No financial bias could be found. The researchers are music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

Sharda 2018

Study characteristics	
Methods	**Allocation:** randomised **Blindness:** assessors blind **Duration:** 8-12 weeks **Design:** parallel group, single centre
Participants	**Diagnosis:** ASD according to the DSM-IV criteria **N:** 51 **Age range:** 6 to 12 years **Sex:** 43 boys, 8 girls **Setting:** unclear **Location:** Montreal, Canada
Interventions	1. **Experimental (n = 26):** MT; use of musical instruments, songs and rhythmic cues while targeting communication, turn-taking, sensorimotor integration, social appropriateness and musical interaction; 45-minute weekly sessions over 8-12 weeks 2. **Control (n = 25):** play-based intervention; to control for non-specific factors, such as positive treatment expectancies, intervention support, therapist attention and emotional engagement; 45-minute weekly sessions over 8-12 weeks
Outcomes	1. **Social interaction**: a. Social Responsiveness Scale (SRS-2); lower scores are favourable; rated by parents pre- and post-intervention b. Joint engagement, measured using a coding scheme for various engagement states observed from session videos; these were not used as group means were only reported in graphs, with no individual participant data available. 2. **Non-verbal communication:** Children's Communication Checklist (CCC-2), measuring pragmatic communication; higher scores are favourable; rated by parents pre- and post-intervention 3. **Verbal communication**: Peabody Picture Vocabulary Test-4 (PPVT-4), measuring language; higher scores are favourable; rated by blinded assessors pre- and post-intervention 4. **Quality of life:** Beach Family Quality of Life (FQoL); higher scores are favourable; rated by parents pre- and post-intervention 5. **Adaptive behaviour:** a. Vineland Adaptive Behaviour Scales (VABS), subscale maladaptive behaviour; lower scores are favourable; rated by parents pre- and post-intervention

Sharda 2018 *(Continued)*

b. Neuroimaging outcomes (rsfMRI: intrinsic brain connectivity of fronto-temporal brain networks); not used in meta-analysis

6. **Movement:** measured using a video-based optical flow analysis method, with whole-body movement amplitude of child and therapist calculated separately; not used since outside of predefined outcome categories

Notes	**Funding source(s):** Canadian Institutes of Health Research; Quebec Bioimaging Network

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Participants were randomized to MT (n = 26) or NM (n = 25) using the covariate-adaptive method where the first 20 participants were randomized using simple coin toss and remaining 31 by the MinimPy software (http://minimpy.sourceforge.net/) by the first author."
Allocation concealment (selection bias)	Low risk	Likely to be concealed. Randomisation was conducted by the first author, who was not involved in assessing behavioural outcomes and was not involved in interventions.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"All other assessors and authors were blind to group allocation information. Our attempt to blind parents (who assessed parent-rated outcomes) was only partially successful, with 31 out of the 51 parents reporting awareness of group allocation."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Registered trial. No deviations from the intended intervention that arose because of the trial context
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy method: yes Adequate music therapy training: yes

Thomas 2003

Study characteristics

Methods	**Allocation:** randomised order of treatment **Blindness:** no blinding **Duration:** 12 weeks **Design:** cross-over (within each session)
Participants	**Diagnosis:** autism **N:** 6 **Age range:** 2 to 3 years

Thomas 2003 *(Continued)*

	Sex: 5 boys, 1 girl **Setting:** unclear **Location:** USA
Interventions	1. **Experimental (n = 6):** MT; using songs, instruments, vocal sounds and movement to interact with the child and musically or verbally respond to the child's verbal or non-verbal behaviour; 12 x 15-minute sessions, immediately following or preceding playtime sessions 2. **Control (n = 6):** playtime; attempts to interact with the child using toys and verbally responding to the child's non-verbal or verbal behaviour; 12 x 15-minute sessions, immediately following or preceding music therapy sessions
Outcomes	Behaviour observation during intervention based on videotaped sessions, coded by trained music therapy intern (inter-rater reliability 0.85) 1. **Social interaction:** requesting behaviour (percentage of session time) 2. **Adaptive behaviour:** on-task behaviour (percentage of session time)
Notes	**Funding source(s):** Mid-Atlantic Region of the American Music Therapy Association

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomised, no further details given
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention was unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details given whether the assessor was blinded to the randomisation result
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts No missing data reported
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researchers are music therapists. Adequate music therapy method: yes Adequate music therapy training: yes

Thompson 2014

Study characteristics

Methods	**Allocation:** randomised

Thompson 2014 *(Continued)*

	Blindness: no blinding **Duration:** 16 weeks (plus 8 weeks of follow-up, but not used, as only for music therapy group) **Design:** parallel group
Participants	**Diagnosis:** ASD **N:** 23 **Age range:** 3 to 6 years **Sex:** 19 boys, 4 girls **Setting:** participants' homes **Location:** Australia
Interventions	1. **Experimental (n = 12):** home-based, family-centred MT (using songs, improvisation, structured music interactions) plus standard care; 16 sessions, scheduled weekly 2. **Control (n = 11):** standard care
Outcomes	1. **Social interaction:** a. Vineland Social Emotional Early Childhood Scales (Vineland SEEC); higher scores are favourable; rated by researcher based on parent's responses at baseline and post-intervention b. Social Responsiveness Scale-Preschool version (SRS-PS); lower scores are favourable; rated by parents pre- and post-intervention 2. **Non-verbal communication:** MacArthur-Bates Communicative Development Inventories-Words and Gestures (MBCDI-W&G), subscale 2 (actions & gestures used); higher scores are favourable; rated by parents pre- and post-intervention 3. **Verbal communication:** MBCDI-W&G, subscales 1B (phrases understood) and 1D (words understood + words produced) assessed pre- and post-intervention 4. **Quality of family relationships:** a. Parent-Child Relationship Inventory (PCRI); higher scores are favourable; rated by parents pre- and post-intervention b. Music Therapy Diagnostic Assessment (MTDA), rated by researcher; not used since rated for music therapy group only
Notes	**Funding source(s):** University of Melbourne; Victorian Department of Education and Early Childhood Development, Australia

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomised (computer-generated random sequence)
Allocation concealment (selection bias)	Low risk	An independent statistician prepared opaque, numbered allocation envelopes.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention is unknown.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Parent-report based measures were used. However, measures contained internal safeguards to address bias as evidenced by high correlations with non-parent rated measures or high test-retest correlations.
Incomplete outcome data (attrition bias)	Low risk	Low dropout rate Intention-to-treat analysis

Thompson 2014 *(Continued)*

All outcomes		
Selective reporting (reporting bias)	Low risk	All outcome measures of interest were considered in the analysis.
Other bias	Low risk	No financial bias could be found. The researcher is a music therapist. Adequate music therapy method: yes Adequate music therapy training: yes

Yurteri 2019

Study characteristics	
Methods	**Allocation:** unclear **Blindness:** not known **Duration:** 8 weeks **Design:** parallel group
Participants	**Diagnosis:** ASD **N:** 24 **Age range:** 4.8 to 9.3 years **Sex:** 24 boys, no girls **Setting:** unclear **Location:** Düzce, Turkey
Interventions	1. **Experimental (n = 12):** improvisational MT; led by a music therapist and "mostly one-on-one, sometimes groups of three"; twice-weekly, 40-minute sessions for 8 weeks 2. **Control (n = 12):** TAU; no treatment except monthly routine child psychiatric follow-up and special education
Outcomes	1. **Quality of life:** Pediatric Quality of Life Inventory (PedsQL) Total score; high scores are favourable; measured pre- and post-intervention 2. **Total autism symptom severity:** Autism Behavior Checklist (AuBC) Total; low scores are favourable; measured pre- and post-intervention
Notes	**Funding source(s):** not reported

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No details given. Probably randomised
Allocation concealment (selection bias)	Unclear risk	No details given
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The fact that children with ASD participating in the study were not blinded was considered unlikely to introduce bias. The possible risk of bias introduced by therapists administering the intervention is unknown.
Blinding of outcome assessment (detection bias)	Unclear risk	Outcome assessments were administered by parents.

Yurteri 2019 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts. No missing data reported
Selective reporting (reporting bias)	Low risk	No trial registration/study protocol. It was likely that there were no deviations from the intended intervention that arose because of the trial context.
Other bias	Low risk	No personal or financial bias could be found. Adequate music therapy training: yes Adequate music therapy method: yes

ABC: Aberrant Behavior Checklist
ADOS(-RRB, -SA): Autism Diagostic Observation Schedule (-Restricted and Repetitive Behaviors, -Social Affect)
ASD: autism spectrum disorder
ASSP: Autism Social Skills Profile
ATEC: Autism Treatment Evaluation Checklist
AuBC: Autism Behavior Checklist
CAMHS: Child and Adolescent Mental Health Services
CARS(-2-HF, -BR): Childhood Autism Rating Scale (-Second Edition-High Functioning, -Brazilian Version)
CBC: Child Behavior Checklist
CCC-2: Children's Communication Checklist-Second Edition
CES-DC: Center for Epidemiological Studies Depression Scale for Children
CGI: Clinical Global Impression scale
DSM: Diagnostic and Statistical Manual of Mental Disorders
DVD: digital video/versatile disc
ECA-R: Revised Clinical Scale for the Evaluation of Autistic Behavior
EQ5D: a term (not an abbreviation) to describe a generic measure of health-related quality of life developed by the EuroQol Group
ESCS: Early Social Communication Scales
FAD: McMaster Family Assessment Device
FQoL: Beach Center Family Quality of Life Scale
ICD: International Classification of Diseases
MACT: music attention control training
MBCDI-W&G: MacArthur-Bates Communicative Development Inventories-Words and Gestures
MPIP: Mother Play Intervention Profile
MT: music therapy
MTDA: Music Therapy Diagnostic Assessment
NM: non-music
PCRI: Parent-Child Relationship Inventory
PDDBI: Pervasive Developmental Disorder Behavior Inventory
PedsQL: Pediatric Quality of Life Inventory
PEP: Psychoeducational Profile
PI: principal investigator
PPVT-4: Peabody Picture Vocabulary Test-4th Edition
QoL: quality of life
rsfMRI: resting-state functional magnetic resonance imaging
SAE: serious adverse effects
SEEC: Vineland Social-Emotional Early Childhood Scales
SRS(-2, -PS): Social Responsiveness Scale (-Second Edition, -Preschool Version)
SSIS: Social Skills Improvement System
SSRS(-P): Social Skills Rating System Scale (-Parent form)
T(0, 1): time
TAU: treatment-as-usual
TEA-Ch 2: Test of Everyday Attention for Children 2
TRIAD: Treatment and Research Institute for Autism Spectrum Disorders
TSC: Trial Steering Committee
TSSA: TRIAD Social Skills Assessment

VABS: Vineland Adaptive Behavior Scales
VPES: Verbal Production Evaluation Scale

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Bringas 2015	Not ASD (severe neurological disorders)
Cowan 2016	Not an RCT or CCT (uncontrolled design)
Dezfoolian 2013	Not an RCT or CCT (uncontrolled design)
Edgerton 1994	Not an RCT or CCT (case series)
Finnigan 2010	Not an RCT or CCT (case study)
Gooding 2011	Not an RCT or CCT (uncontrolled design)
Hairston 1990	Not an RCT or CCT (case series)
Iseri 2014	Not an RCT or CCT (uncontrolled design)
Kern 2006	Not an RCT or CCT (case series)
Kern 2007	Not an RCT or CCT (case study)
Kim 2000	Not an RCT or CCT (uncontrolled design)
Mendelson 2016	No relevant comparison condition (both groups MT)
Sanglakh Goochan Atigh 2017	Not MT (movement activities with music)
Saperston 1973	Not an RCT or CCT (case study)
Thaut 1988	Not an intervention study (assessment)
Yoo 2018	Not an RCT or CCT (uncontrolled design)

ASD: autism spectrum disorder
CCT: controlled clinical trial
MT: music therapy
RCT: randomised controlled trial

Characteristics of studies awaiting classification *[ordered by study ID]*

NCT03267095

Methods	**Allocation:** randomised **Blindness:** none **Duration:** 12 months **Design:** parallel group
Participants	**Diagnosis:** autism spectrum disorder; IQ ≥ 75 **Estimated N:** 60 **Ages eligible:** 3 to 7 years **Sexes eligible:** all

NCT03267095 *(Continued)*	
	Setting: unknown **Location:** Egypt
Interventions	1. Experimental: **music therapy** sessions for 12 months; songs composed by the therapist and target words (selected from functional vocabularies that typically developing 3-year-old children can use effectively in everyday interactions) will be used for the study 2. Control:**counselling** (setting, duration, and number of sessions unclear)
Outcomes	1. Arabic language test, pre- and post intervention
Notes	**Comment:** awaiting classification as the trial registration does not contain sufficient information on some aspects of the design (randomisation, diagnosis, types of interventions) to ensure all eligibility criteria are met; recruitment has not started **Funding source:** Assiut University

IQ: Intelligence quotient
N: number

Characteristics of ongoing studies *[ordered by study ID]*

ISRCTN18340173	
Study name	**Public title:** Improvisational music therapy for children with autism spectrum disorder assessed using brain imaging **Scientific title:** Music therapy outcome study for children with Autism Spectrum Disorder through integration of child neuroimaging and neuropsychology: an exploratory study
Methods	**Allocation:** non-randomised (propensity score matching) **Blindness:** single (outcomes assessor) **Duration:** 1 year **Design:** parallel
Participants	**Diagnosis:** children with ASD (confirmed by DSM-5 and ADOS) **Exclusion criteria:** hearing/visual impairment, congenital genetic disease (cerebral palsy); seizure and other neurological condition; other psychiatric disorder such as ADHD, schizophrenia, depression, bi-polar; previous and on-going experiences of music therapy; medication such as risperidone, aripirazole, SSRI, etc. **Estimated N:** 50 **Ages eligible:** 24-72 months **Sexes eligible:** all **Setting:** Jeonju University, Seoul National University Hospital, and four music therapy centres **Location:** Seoul, Jeonju, and Namyangju-si, South Korea
Interventions	1. **Experimental:** improvisational music therapy, once weekly, for one year 2. **Control:** standard care
Outcomes	**Primary outcome measures** 1. Severity of autism measured using the Autism Diagnostic Observation Schedule (ADOS), pre-, mid-(6 months), and post-intervention 2. Severity of autism measured using the Childhood Autism Rating Scale-2 (CARS-2),pre-, mid-(6 months), and post-intervention **Secondary outcome measures** Collected pre-, mid-(6 months), and post-intervention: 1. Social Responsiveness Scale (SRS)

ISRCTN18340173 *(Continued)*

	2. Aberrant Behavior Checklist (ABC) 3. Social Communication Questionnaire (SCQ) 4. Social Maturity Scale (SMS) 5. Child Behavior Checklist 1.5-5 (CBCL 1.5-5) 6. Beck Depression Inventory (BDI) 7. State-Trait Anxiety Inventory (STAI-T/S) 8. Maternal Behavior Research Instrument (MBRI) 9. Korean Parenting Stress Index (PSI) 10.Mental Health Continuum-Short Form (MHC-SF) 11.Carer-QoL-7D/Child QoL (VAS) 12.Concomitant treatment form Collected pre- and post-intervention: 1. Neuroimaging: aMRI/DTI/rsfMRI 2. Psychoeducational Profile-Revised (PEP-R) Baseline test only: Blood and Urine for detecting DNA and harmful environmental exposure Post-trial in-depth interview with the mothers and the therapist (thematic analysis; qualitative study)
Starting date	1 August 2018
Contact information	Jinah Kim, Professor, Jeonju University, Jeonju, South Korea
Notes	**Funding source:** National Research Foundation of Korea

NCT03560297

Study name	**Public title:** SeRenade parent-child music class program **Scientific title:** SeRenade parent-child music class program for families of children with and without ASD
Methods	**Allocation:** randomised **Blindness:** single (outcomes assessor) **Duration:** 12 weeks **Design:** cross-over
Participants	**Diagnosis:** children diagnosed with ASD or children without ASD **Exclusion criteria:** currently enrolled in music therapy or parent-child music classes (for children with ASD); significant behaviour problems (e.g. aggression toward other children); significant hearing or visual impairments **Estimated N:** 68 **Ages eligible:** 20-72 months **Sexes eligible:** all **Setting:** Vanderbilt University Medical Center **Location:** Nashville, Tennessee, USA
Interventions	1. **Experimental:** SeRenade programme: parent-child music class programme (parent training, peer inclusion, musical play) 2. **Control:** delayed/waiting-list programme: participants do not participate in the programme for a time period
Outcomes	**Primary outcome measures**

NCT03560297 *(Continued)*

	1. Motor Imitation Scale (standardised, elicited imitation assessment; scores range 0-32; greater scores indicate greater imitation skills), pre- and post-intervention 2. Actions & Gestures (standardised, parent-report questionnaire of child's gestures/actions/pretend play), pre- and post-intervention 3. Parenting Stress Index, 4th edition Short Form (standardised, parent-report questionnaire), pre- and post-intervention 4. Parenting efficacy/quality (parent-report questionnaire; range 10-40; higher scores indicate greater efficacy), pre- and post-intervention **Secondary outcome measures** 1. Parent-child interaction (parent-child play/music session), pre- and post-intervention 2. Social visual attention eye-tracking paradigm, pre- and post-intervention **Other outcome measures** 1. Parent mood/connection (parent ratings), pre- and post-intervention 2. Programme acceptability (parent programme evaluation survey), post-intervention
Starting date	15 May 2018
Contact information	Miriam Lense, Research Instructor, Vanderbilt University Medical Center, Nashville, Tennessee, USA, 37232
Notes	**Funding sources:** Vanderbilt University Medical Center; National Endowment for the Arts; USA

NCT04557488

Study name	**Public title:** Effectiveness of music therapy in social skill intervention for children with ASD/ID **Scientific title:** Effectiveness of music therapy in social skill intervention for children with ASD/ID: a randomized controlled trial
Methods	**Allocation:** randomised **Blindness:** single (outcomes assessor) **Duration:** 12 weeks **Design:** parallel
Participants	**Diagnosis:** children with a formal clinical diagnosis of ASD and an overall and verbal IQ of 50-84 as assessed by certified clinician **Exclusion criteria:** severe physical or sensory disabilities (e.g. deafness); other neurodevelopmental, psychiatric, or neurological comorbidities or on prescribed psychiatric medication **Estimated N:** 80 **Ages eligible:** 6-13 years **Sexes eligible:** all **Setting:** The Education University of Hong Kong **Location:** Hong Kong
Interventions	1. **Experimental:** Group music therapy: sessions follow a similar structure (hello song, musical activities, goodbye song); musical activities will vary in each session and will be mixed in later sessions to revisit and practice social skills; groups of eight, led by certified music therapist 2. **Control:** Behavioural-based social skill group training: sessions follow a standard structure (opening greetings, social activities according to the theme of the session, closing activity); activities and games will vary in each session and will be mixed in later session to revisit and practice social skills; groups of eight, led by registered social worker with experience in providing social skill training for children with ASD and ID
Outcomes	**Primary outcome measures**

NCT04557488 *(Continued)*

	1. Childhood Autism Rating Scale-2; pre- and post-intervention and follow-up 4 months after intervention 2. Social Responsiveness Scale second edition; pre- and post-intervention and follow-up 4 months after intervention 3. In-session social behaviour (coding of videos); first and last session **Secondary outcome measures** 1. EEG recordings (in three conditions: resting state, social scenes, preferred music; for 5 minutes each); pre- and post-intervention and follow-up 4 months after intervention
Starting date	1 October 2020
Contact information	Yen Na Cherry Yum, Assistant Professor, Education University of Hong Kong, Hong Kong
Notes	**Funding source:** Education University of Hong Kong

NCT04936048

Study name	**Public title:** Music for autism (M4A) **Scientific title:** Music for autism: binational randomised controlled trial of music therapy versus play therapy for autistic children
Methods	**Allocation:** randomised **Blindness:** single (outcomes assessor) **Duration:** 12 weeks **Design:** cross-over
Participants	**Diagnosis:** children diagnosed with ASD by a licensed clinical professional using standardised diagnostic tools (ADOS, ADI-R) **Exclusion criteria:** recent or current music therapy; metallic or electronic implants **Estimated N:** 80 **Ages eligible:** 6-12 years **Sexes eligible:** all **Setting:** University of Vienna, Vienna, Austria; NORCE Norwegian Research Centre, Bergen, Norway **Locations:** Vienna, Austria; Bergen, Norway
Interventions	1. **Experimental:** music therapy (MT) using rhythmic cues, music instruments, songs, and stories accompanied by songs or musical instruments to target common goals (multisensory integration, verbal and social communication, emotion regulation, turn-taking, social appropriateness, interaction); 12 weekly one-on-one sessions of 45 minutes each, conducted by a licensed music therapist 2. **Control:** play therapy using verbal interaction, toys (Lego, finger puppets, Play Doh, puzzles), and the same stories as in MT, but without a musical component, to target the same common goals as MT; 12 weekly one-on-one sessions of 45 minutes each, conducted by a licensed music therapist
Outcomes	**Primary outcome measures** 1. Children's Communication Checklist-2; pre- and post-intervention 2. Brain connectivity of frontotemporal regions; pre- and post-intervention **Secondary outcome measures** 1. Child and Adolescent Scale of Participation; pre- and post-intervention 2. Beach Center Family Quality of Life Scale; pre- and post-intervention 3. Peabody Picture Vocabulary Test-4th edition; pre- and post-intervention 4. Social Responsiveness Scale; pre- and post-intervention 5. Vineland Adaptive Behavior Scales; pre- and post-intervention

NCT04936048 *(Continued)*	6. Hair cortisol concentration; pre- and post-intervention 7. Grey and white matter volume (structural brain changes); pre- and post-intervention
Starting date	1 August 2021
Contact information	Christian Gold, NORCE Norwegian Research Centre AS, Bergen, Norway
Notes	**Funding sources:** NORCE Norwegian Research Centre AS & University of Bergen, Norway; University of Vienna, Austria

ABC: Aberrant Behavior Checklist
ADHD: attention deficit hyperactivity disorder
ADI-R: Autism Diagnostic Interview-Revised
ADOS: Autism Diagostic Observation Schedule
aMRI/DTI/rsfMRI: advanced magnetic resonance imaging/diffusion tensor imaging/resting-state functional magnetic resonance imaging
ASD: autism spectrum disorder
BDI: Beck Depression Inventory
CARS: Childhood Autism Rating Scale
CBCL: Child Behavior Checklist
DNA: deoxyribonucleic acid
DSM: Diagnostic and Statistical Manual of Mental Disorders
EEG: electroencephalogram
ID: intellectual disability
IQ: intelligence quotient
M4A: Music for Autism
MBRI: Maternal Behavior Research Instrument
MHC-SF: Mental Health Continuum-Short Form
MT: music therapy
PEP-R: Psychoeducational Profile Revised
PSI: Parenting Stress Index
QoL: quality of life
SCQ: Social Communication Questionnaire
SMS: Social Maturity Scale
SRS: Social Responsiveness Scale
SSRI: selective serotonin reuptake inhibitors
STAI-T/S: State-Trait Anxiety Inventory
VAS: visual analogue scale

DATA AND ANALYSES

Comparison 1. Music therapy vs placebo therapy or standard care

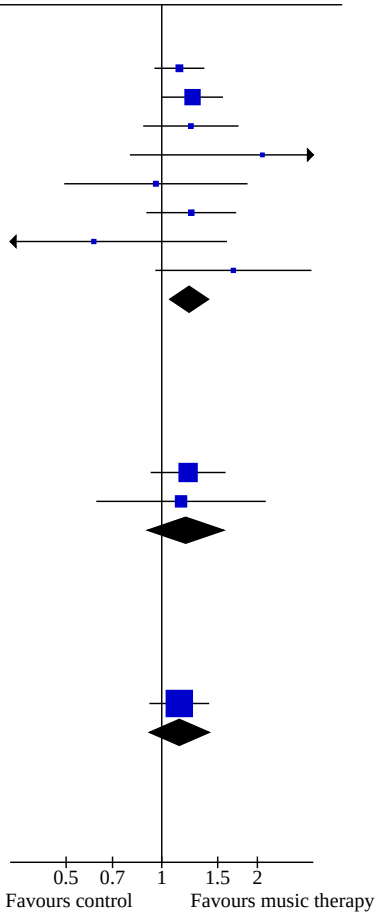
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Global improvement	8		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1.1 Immediately post-intervention	8	583	Risk Ratio (M-H, Fixed, 95% CI)	1.22 [1.06, 1.40]
1.1.2 1-5 months post-intervention	2	99	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [0.90, 1.57]
1.1.3 6-11 months post-intervention	1	364	Risk Ratio (M-H, Fixed, 95% CI)	1.14 [0.91, 1.41]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.2 Social interaction	14		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.2.1 During intervention	3	44	Std. Mean Difference (IV, Random, 95% CI)	1.15 [0.49, 1.80]
1.2.2 Immediately post-intervention	12	603	Std. Mean Difference (IV, Random, 95% CI)	0.26 [-0.05, 0.57]
1.2.3 1-5 months post-intervention	2	59	Std. Mean Difference (IV, Random, 95% CI)	0.54 [-0.11, 1.19]
1.2.4 6-11 months post-intervention	1	258	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.30, 0.18]
1.3 Non-verbal communication	9		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.3.1 During intervention	3	50	Std. Mean Difference (IV, Fixed, 95% CI)	1.06 [0.44, 1.69]
1.3.2 Immediately post-intervention	7	192	Std. Mean Difference (IV, Fixed, 95% CI)	0.26 [-0.03, 0.55]
1.4 Verbal communication	12		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.4.1 During intervention	4	129	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.41, 0.28]
1.4.2 Immediately post-intervention	8	276	Std. Mean Difference (IV, Random, 95% CI)	0.30 [-0.18, 0.78]
1.4.3 1-5 months post-intervention	1	52	Std. Mean Difference (IV, Random, 95% CI)	0.22 [-0.33, 0.76]
1.5 Quality of life	3		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.5.1 Immediately post-intervention	3	340	Std. Mean Difference (IV, Fixed, 95% CI)	0.28 [0.06, 0.49]
1.5.2 6-11 months post-intervention	1	249	Std. Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.21, 0.29]
1.6 Total autism symptom severity	9		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.6.1 During intervention	1	16	Std. Mean Difference (IV, Random, 95% CI)	0.15 [-0.83, 1.14]
1.6.2 Immediately post-intervention	9	575	Std. Mean Difference (IV, Random, 95% CI)	-0.83 [-1.41, -0.24]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.6.3 1-5 months post-intervention	2	69	Std. Mean Difference (IV, Random, 95% CI)	-0.93 [-1.81, -0.06]
1.6.4 6-11 months post-intervention	1	289	Std. Mean Difference (IV, Random, 95% CI)	0.18 [-0.05, 0.41]
1.7 Adverse events	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.7.1 Immediately post-intervention	1	290	Risk Ratio (M-H, Fixed, 95% CI)	1.52 [0.39, 5.94]
1.7.2 6-11 months post-intervention	1	290	Risk Ratio (M-H, Fixed, 95% CI)	0.88 [0.23, 3.46]
1.8 Adaptive behaviour	9		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.8.1 During intervention	4	52	Std. Mean Difference (IV, Fixed, 95% CI)	1.19 [0.56, 1.82]
1.8.2 Immediately post-intervention	5	462	Std. Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.20, 0.16]
1.8.3 1-5 months post-intervention	1	35	Std. Mean Difference (IV, Fixed, 95% CI)	0.56 [-0.12, 1.24]
1.8.4 6-11 months post-intervention	1	290	Std. Mean Difference (IV, Fixed, 95% CI)	-0.12 [-0.36, 0.11]
1.9 Quality of family relationships	3		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.9.1 Immediately post-intervention	3	56	Std. Mean Difference (IV, Fixed, 95% CI)	0.29 [-0.24, 0.83]
1.9.2 1-5 months post-intervention	1	15	Std. Mean Difference (IV, Fixed, 95% CI)	-0.04 [-1.07, 0.99]
1.10 Identity formation	2		Std. Mean Difference (IV, Random, 95% CI)	Subtotals only
1.10.1 Immediately post-intervention	2	55	Std. Mean Difference (IV, Random, 95% CI)	1.35 [-0.58, 3.28]
1.10.2 1-5 months post-intervention	1	35	Std. Mean Difference (IV, Random, 95% CI)	0.86 [0.16, 1.55]
1.11 Depression	1		Std. Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.11.1 Immediately post-intervention	1	34	Std. Mean Difference (IV, Fixed, 95% CI)	-0.34 [-1.01, 0.34]
1.11.2 1-5 months post-intervention	1	36	Std. Mean Difference (IV, Fixed, 95% CI)	-0.60 [-1.27, 0.07]

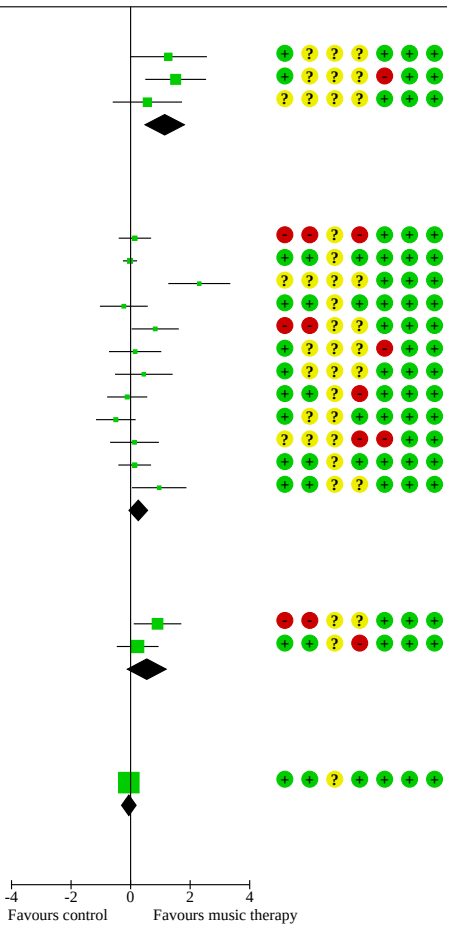
Analysis 1.1. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 1: Global improvement

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.1.1 Immediately post-intervention						
Bharathi 2019	25	26	22	26	15.1%	1.14 [0.95 , 1.36]
Bieleninik 2017	95	182	76	182	52.1%	1.25 [1.00 , 1.56]
Kim 2008	10	10	8	10	5.8%	1.24 [0.87 , 1.75]
LaGasse 2014	7	9	3	8	2.2%	2.07 [0.79 , 5.42]
Porter 2017	10	24	10	23	7.0%	0.96 [0.49 , 1.86]
Rabeyron 2020	17	19	13	18	9.2%	1.24 [0.89 , 1.72]
Schwartzberg 2013	4	12	6	11	4.3%	0.61 [0.23 , 1.60]
Thompson 2014	11	12	6	11	4.3%	1.68 [0.95 , 2.96]
Subtotal (95% CI)		**294**		**289**	**100.0%**	**1.22 [1.06 , 1.40]**
Total events:	179		144			
Heterogeneity: Chi² = 5.53, df = 7 (P = 0.60); I² = 0%						
Test for overall effect: Z = 2.77 (P = 0.006)						
1.1.2 1-5 months post-intervention						
Bharathi 2019	23	26	19	26	65.0%	1.21 [0.92 , 1.59]
Porter 2017	12	24	10	23	35.0%	1.15 [0.62 , 2.13]
Subtotal (95% CI)		**50**		**49**	**100.0%**	**1.19 [0.90 , 1.57]**
Total events:	35		29			
Heterogeneity: Chi² = 0.03, df = 1 (P = 0.87); I² = 0%						
Test for overall effect: Z = 1.23 (P = 0.22)						
1.1.3 6-11 months post-intervention						
Bieleninik 2017	92	182	81	182	100.0%	1.14 [0.91 , 1.41]
Subtotal (95% CI)		**182**		**182**	**100.0%**	**1.14 [0.91 , 1.41]**
Total events:	92		81			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.15 (P = 0.25)						
Test for subgroup differences: Chi² = 0.29, df = 2 (P = 0.86), I² = 0%						



Analysis 1.2. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 2: Social interaction

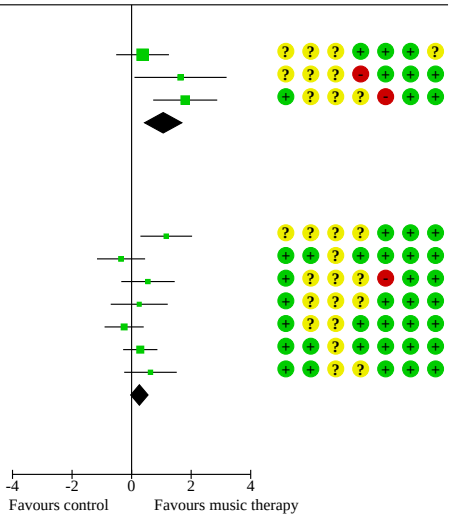
Study or Subgroup	MT Mean	MT SD	MT Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.2.1 During intervention								
Arezina 2011	10.87	2.29	6	7.2	2.99	6	26.1%	1.27 [-0.02 , 2.56]
Kim 2008	1.58	1	10	0	1	10	41.8%	1.51 [0.49 , 2.53]
Thomas 2003	0.61	1	6	0	1	6	32.1%	0.56 [-0.60 , 1.73]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**1.15 [0.49 , 1.80]**
Heterogeneity: Tau² = 0.00; Chi² = 1.50, df = 2 (P = 0.47); I² = 0%								
Test for overall effect: Z = 3.40 (P = 0.0007)								
1.2.2 Immediately post-intervention								
Bharathi 2019	114.85	15.07	26	112.62	15.27	26	10.3%	0.14 [-0.40 , 0.69]
Bieleninik 2017	-0.02	1	142	0	1	129	13.8%	-0.02 [-0.26 , 0.22]
Chen 2013	-10.67	3	18	-17.81	3	9	5.6%	2.31 [1.27 , 3.35]
Gattino 2011	-12.25	1.54	12	-11.92	1.24	12	7.5%	-0.23 [-1.03 , 0.58]
Ghasemtabar 2015	30.55	4	13	27.34	3.54	14	7.6%	0.83 [0.03 , 1.62]
Kim 2008	50.5	17.54	10	47.6	18.9	10	6.9%	0.15 [-0.73 , 1.03]
LaGasse 2014	-93.5	17.57	9	-103.88	26.52	8	6.2%	0.44 [-0.52 , 1.41]
Porter 2017	81.88	12.82	16	83.42	13.34	18	8.8%	-0.11 [-0.79 , 0.56]
Rabeyron 2020	-12.95	3.34	19	-11.09	4	17	8.9%	-0.50 [-1.16 , 0.17]
Schwartzberg 2013	97.56	13	12	95.79	13.06	11	7.4%	0.13 [-0.69 , 0.95]
Sharda 2018	-69.36	10.89	26	-70.8	9.66	25	10.3%	0.14 [-0.41 , 0.69]
Thompson 2014	1	1	11	0	1	10	6.6%	0.96 [0.05 , 1.87]
Subtotal (95% CI)			**314**			**289**	**100.0%**	**0.26 [-0.05 , 0.57]**
Heterogeneity: Tau² = 0.17; Chi² = 29.51, df = 11 (P = 0.002); I² = 63%								
Test for overall effect: Z = 1.62 (P = 0.11)								
1.2.3 1-5 months post-intervention								
Ghasemtabar 2015	30.61	4.25	13	26.85	3.82	14	45.7%	0.90 [0.11 , 1.70]
Porter 2017	82.57	13.96	14	79.14	14.14	18	54.3%	0.24 [-0.46 , 0.94]
Subtotal (95% CI)			**27**			**32**	**100.0%**	**0.54 [-0.11 , 1.19]**
Heterogeneity: Tau² = 0.08; Chi² = 1.51, df = 1 (P = 0.22); I² = 34%								
Test for overall effect: Z = 1.63 (P = 0.10)								
1.2.4 6-11 months post-intervention								
Bieleninik 2017	-0.06	1	132	0	1	126	100.0%	-0.06 [-0.30 , 0.18]
Subtotal (95% CI)			**132**			**126**	**100.0%**	**-0.06 [-0.30 , 0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.48 (P = 0.63)								
Test for subgroup differences: Chi² = 13.36, df = 3 (P = 0.004), I² = 77.5%								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.3. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 3: Non-verbal communication

Study or Subgroup	MT Mean	MT SD	MT Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F G
1.3.1 During intervention									
Buday 1995	5.1	2.89	10	4	2.83	10	49.8%	0.37 [-0.52 , 1.25]	? ? ? + + + ?
Farmer 2003	28.85	15.69	5	8.05	3.9	5	16.3%	1.64 [0.10 , 3.19]	? ? ? - + + +
Kim 2008	1.88	1	10	0	1	10	33.8%	1.80 [0.73 , 2.88]	+ ? ? ? - + +
Subtotal (95% CI)			**25**			**25**	**100.0%**	**1.06 [0.44 , 1.69]**	
Heterogeneity: Chi² = 4.71, df = 2 (P = 0.09); I² = 58%									
Test for overall effect: Z = 3.33 (P = 0.0009)									
1.3.2 Immediately post-intervention									
Chen 2013	-1.83	0.7	18	-2.67	0.7	9	11.2%	1.16 [0.30 , 2.03]	? ? ? ? + + +
Gattino 2011	-2.5	0.37	12	-2.33	0.54	12	13.0%	-0.35 [-1.16 , 0.45]	+ + ? + + + +
Kim 2008	29.1	15.33	10	20.2	15.64	10	10.5%	0.55 [-0.35 , 1.45]	+ ? ? ? - + +
LaGasse 2014	0.27	1	9	0	1	8	9.2%	0.26 [-0.70 , 1.21]	+ ? ? ? + + +
Rabeyron 2020	-2.13	0.81	19	-1.91	0.91	17	19.6%	-0.25 [-0.91 , 0.41]	+ ? ? + + + +
Sharda 2018	80.46	15.26	24	76.43	11.64	23	25.6%	0.29 [-0.28 , 0.87]	+ + ? + + + +
Thompson 2014	34.91	10.83	11	27.8	10.7	10	10.9%	0.63 [-0.25 , 1.52]	+ + ? ? + + +
Subtotal (95% CI)			**103**			**89**	**100.0%**	**0.26 [-0.03 , 0.55]**	
Heterogeneity: Chi² = 9.81, df = 6 (P = 0.13); I² = 39%									
Test for overall effect: Z = 1.76 (P = 0.08)									
Test for subgroup differences: Chi² = 5.18, df = 1 (P = 0.02), I² = 80.7%									

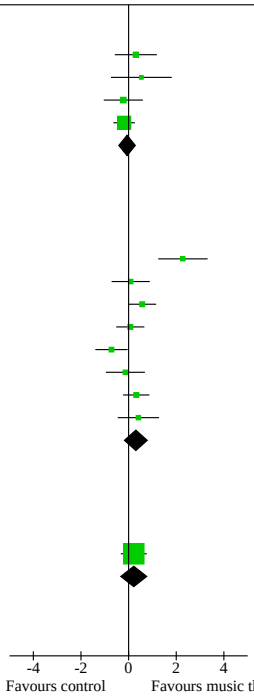


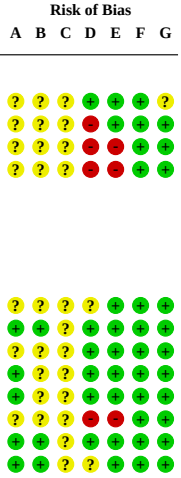
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.4. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 4: Verbal communication

Study or Subgroup	MT Mean	MT SD	MT Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.4.1 During intervention								
Buday 1995	4.2	3.36	10	3.2	2.94	10	15.5%	0.30 [-0.58 , 1.19]
Farmer 2003	113.3	129.8	5	53.15	60.85	5	7.4%	0.54 [-0.74 , 1.81]
Schwartzberg 2013	3.22	1.56	12	3.54	1.17	11	17.9%	-0.22 [-1.04 , 0.60]
Schwartzberg 2016	3.86	1.06	38	4.05	0.93	38	59.3%	-0.19 [-0.64 , 0.26]
Subtotal (95% CI)			**65**			**64**	**100.0%**	**-0.06 [-0.41 , 0.28]**
Heterogeneity: Tau² = 0.00; Chi² = 1.95, df = 3 (P = 0.58); I² = 0%								
Test for overall effect: Z = 0.37 (P = 0.71)								
1.4.2 Immediately post-intervention								
Chen 2013	-2.33	0.71	18	-4	0.71	9	9.7%	2.28 [1.24 , 3.32]
Gattino 2011	-2.54	0.45	12	-2.58	0.44	12	11.9%	0.09 [-0.71 , 0.89]
Lim 2010	133.11	77.23	18	89.62	74.06	32	14.0%	0.57 [-0.02 , 1.16]
Lim 2011	5.39	75	22	0	75	22	14.0%	0.07 [-0.52 , 0.66]
Rabeyron 2020	-2.92	0.92	19	-2.26	0.87	17	13.1%	-0.72 [-1.40 , -0.04]
Schwartzberg 2013	4.12	1.26	12	4.29	1.14	11	11.7%	-0.14 [-0.96 , 0.68]
Sharda 2018	95.04	26.44	26	85.92	29.4	25	14.4%	0.32 [-0.23 , 0.87]
Thompson 2014	232.82	147.42	11	178.7	98.18	10	11.2%	0.41 [-0.46 , 1.28]
Subtotal (95% CI)			**138**			**138**	**100.0%**	**0.30 [-0.18 , 0.78]**
Heterogeneity: Tau² = 0.34; Chi² = 25.30, df = 7 (P = 0.0007); I² = 72%								
Test for overall effect: Z = 1.24 (P = 0.21)								
1.4.3 1-5 months post-intervention								
Bharathi 2019	82.46	18.73	26	78.73	14.99	26	100.0%	0.22 [-0.33 , 0.76]
Subtotal (95% CI)			**26**			**26**	**100.0%**	**0.22 [-0.33 , 0.76]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.78 (P = 0.44)								
Test for subgroup differences: Chi² = 1.73, df = 2 (P = 0.42), I² = 0%								

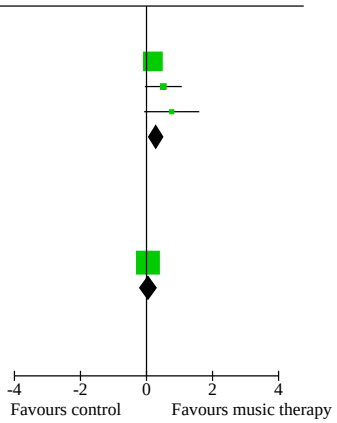




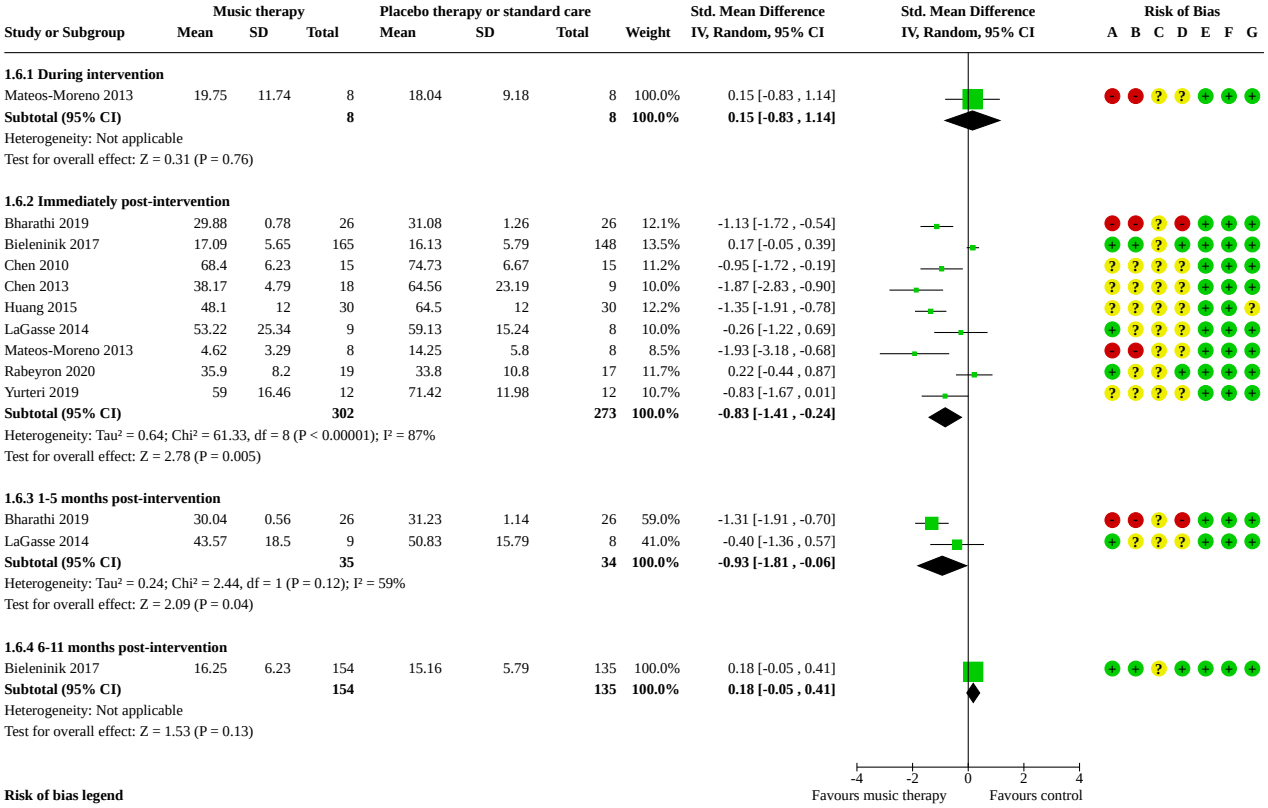
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Analysis 1.5. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 5: Quality of life

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
1.5.1 Immediately post-intervention								
Bieleninik 2017	75.76	14.4	136	72.76	16.29	129	78.7%	0.19 [-0.05 , 0.44]
Sharda 2018	105.36	9.58	26	99.96	11.29	25	14.7%	0.51 [-0.05 , 1.07]
Yurteri 2019	58.88	9.33	12	52.4	6.86	12	6.6%	0.76 [-0.07 , 1.60]
Subtotal (95% CI)			**174**			**166**	**100.0%**	**0.28 [0.06 , 0.49]**
Heterogeneity: Chi² = 2.41, df = 2 (P = 0.30); I² = 17%								
Test for overall effect: Z = 2.55 (P = 0.01)								
1.5.2 6-11 months post-intervention								
Bieleninik 2017	75.16	16.08	125	74.48	15.27	124	100.0%	0.04 [-0.21 , 0.29]
Subtotal (95% CI)			**125**			**124**	**100.0%**	**0.04 [-0.21 , 0.29]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.34 (P = 0.73)								
Test for subgroup differences: Chi² = 1.98, df = 1 (P = 0.16), I² = 49.4%								



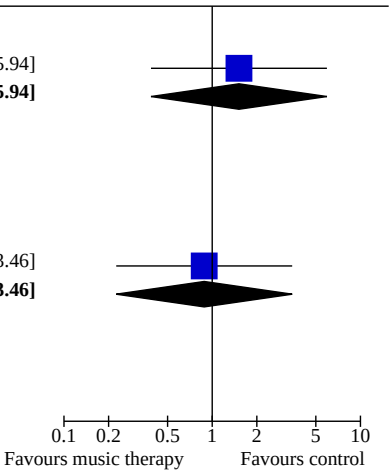
Analysis 1.6. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 6: Total autism symptom severity



Study or Subgroup	Music therapy Mean	SD	Total	Placebo therapy or standard care Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
1.6.1 During intervention									
Mateos-Moreno 2013	19.75	11.74	8	18.04	9.18	8	100.0%	0.15 [-0.83 , 1.14]	
Subtotal (95% CI)			**8**			**8**	**100.0%**	**0.15 [-0.83 , 1.14]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.31 (P = 0.76)									
1.6.2 Immediately post-intervention									
Bharathi 2019	29.88	0.78	26	31.08	1.26	26	12.1%	-1.13 [-1.72 , -0.54]	
Bieleninik 2017	17.09	5.65	165	16.13	5.79	148	13.5%	0.17 [-0.05 , 0.39]	
Chen 2010	68.4	6.23	15	74.73	6.67	15	11.2%	-0.95 [-1.72 , -0.19]	
Chen 2013	38.17	4.79	18	64.56	23.19	9	10.0%	-1.87 [-2.83 , -0.90]	
Huang 2015	48.1	12	30	64.5	12	30	12.2%	-1.35 [-1.91 , -0.78]	
LaGasse 2014	53.22	25.34	9	59.13	15.24	8	10.0%	-0.26 [-1.22 , 0.69]	
Mateos-Moreno 2013	4.62	3.29	8	14.25	5.8	8	8.5%	-1.93 [-3.18 , -0.68]	
Rabeyron 2020	35.9	8.2	19	33.8	10.8	17	11.7%	0.22 [-0.44 , 0.87]	
Yurteri 2019	59	16.46	12	71.42	11.98	12	10.7%	-0.83 [-1.67 , 0.01]	
Subtotal (95% CI)			**302**			**273**	**100.0%**	**-0.83 [-1.41 , -0.24]**	
Heterogeneity: Tau² = 0.64; Chi² = 61.33, df = 8 (P < 0.00001); I² = 87%									
Test for overall effect: Z = 2.78 (P = 0.005)									
1.6.3 1-5 months post-intervention									
Bharathi 2019	30.04	0.56	26	31.23	1.14	26	59.0%	-1.31 [-1.91 , -0.70]	
LaGasse 2014	43.57	18.5	9	50.83	15.79	8	41.0%	-0.40 [-1.36 , 0.57]	
Subtotal (95% CI)			**35**			**34**	**100.0%**	**-0.93 [-1.81 , -0.06]**	
Heterogeneity: Tau² = 0.24; Chi² = 2.44, df = 1 (P = 0.12); I² = 59%									
Test for overall effect: Z = 2.09 (P = 0.04)									
1.6.4 6-11 months post-intervention									
Bieleninik 2017	16.25	6.23	154	15.16	5.79	135	100.0%	0.18 [-0.05 , 0.41]	
Subtotal (95% CI)			**154**			**135**	**100.0%**	**0.18 [-0.05 , 0.41]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.53 (P = 0.13)									

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

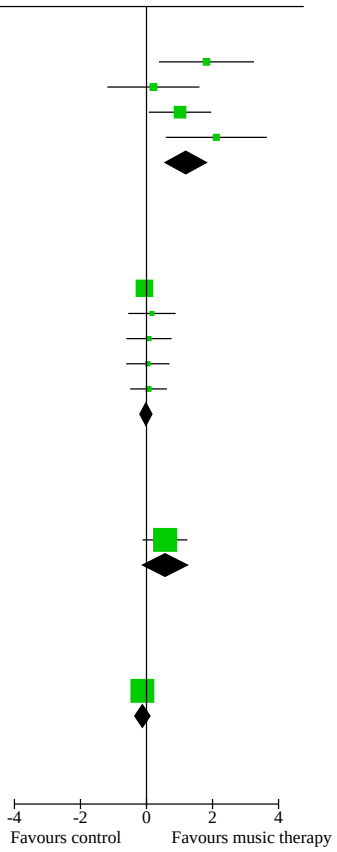
Analysis 1.7. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 7: Adverse events



Study or Subgroup	Music therapy Events	Total	Placebo or standard care Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.7.1 Immediately post-intervention						
Bieleninik 2017	6	165	3	125	100.0%	1.52 [0.39 , 5.94]
Subtotal (95% CI)		**165**		**125**	**100.0%**	**1.52 [0.39 , 5.94]**
Total events:	6		3			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.60 (P = 0.55)						
1.7.2 6-11 months post-intervention						
Bieleninik 2017	4	154	4	136	100.0%	0.88 [0.23 , 3.46]
Subtotal (95% CI)		**154**		**136**	**100.0%**	**0.88 [0.23 , 3.46]**
Total events:	4		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.18 (P = 0.86)						
Test for subgroup differences: Chi² = 0.30, df = 1 (P = 0.58), I² = 0%						

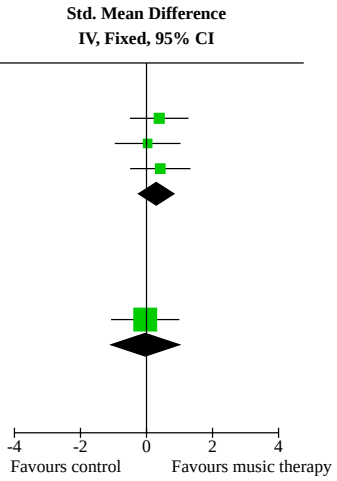
Analysis 1.8. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 8: Adaptive behaviour

Study or Subgroup	MT Mean	MT SD	MT Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
1.8.1 During intervention								
Arezina 2011	89.2	12.18	6	51.08	24.53	6	19.0%	1.82 [0.38 , 3.25]
Brownell 2002	-10.62	8.53	4	-12.5	6.99	4	20.2%	0.21 [-1.18 , 1.60]
Kim 2008	1.06	1	10	0	1	10	44.0%	1.02 [0.07 , 1.96]
Thomas 2003	632.67	115.15	6	407.17	78.35	6	16.7%	2.11 [0.58 , 3.64]
Subtotal (95% CI)			**26**			**26**	**100.0%**	**1.19 [0.56 , 1.82]**
Heterogeneity: Chi² = 4.16, df = 3 (P = 0.24); I² = 28%								
Test for overall effect: Z = 3.72 (P = 0.0002)								
1.8.2 Immediately post-intervention								
Bieleninik 2017	-3.82	1.94	165	-3.69	1.99	148	67.8%	-0.07 [-0.29 , 0.16]
Chen 2010	55.47	15.68	15	52.27	21.46	15	6.5%	0.17 [-0.55 , 0.88]
Porter 2017	-84.53	33.01	15	-87	30.97	18	7.1%	0.08 [-0.61 , 0.76]
Rabeyron 2020	-48.6	27.2	19	-49.8	28.6	17	7.8%	0.04 [-0.61 , 0.70]
Sharda 2018	-19.42	1.76	26	-19.54	2.04	24	10.8%	0.06 [-0.49 , 0.62]
Subtotal (95% CI)			**240**			**222**	**100.0%**	**-0.02 [-0.20 , 0.16]**
Heterogeneity: Chi² = 0.62, df = 4 (P = 0.96); I² = 0%								
Test for overall effect: Z = 0.20 (P = 0.84)								
1.8.3 1-5 months post-intervention								
Porter 2017	-71.19	33.43	16	-89.55	30.8	19	100.0%	0.56 [-0.12 , 1.24]
Subtotal (95% CI)			**16**			**19**	**100.0%**	**0.56 [-0.12 , 1.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.62 (P = 0.11)								
1.8.4 6-11 months post-intervention								
Bieleninik 2017	-3.66	2.11	155	-3.4	2.05	135	100.0%	-0.12 [-0.36 , 0.11]
Subtotal (95% CI)			**155**			**135**	**100.0%**	**-0.12 [-0.36 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.06 (P = 0.29)								
Test for subgroup differences: Chi² = 17.46, df = 3 (P = 0.0006), I² = 82.8%								

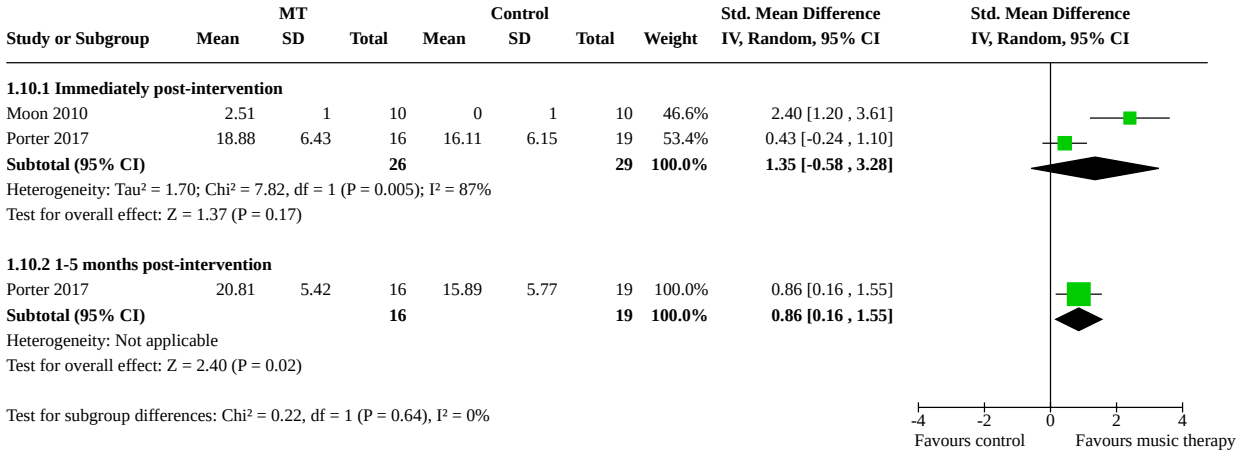


Analysis 1.9. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 9: Quality of family relationships

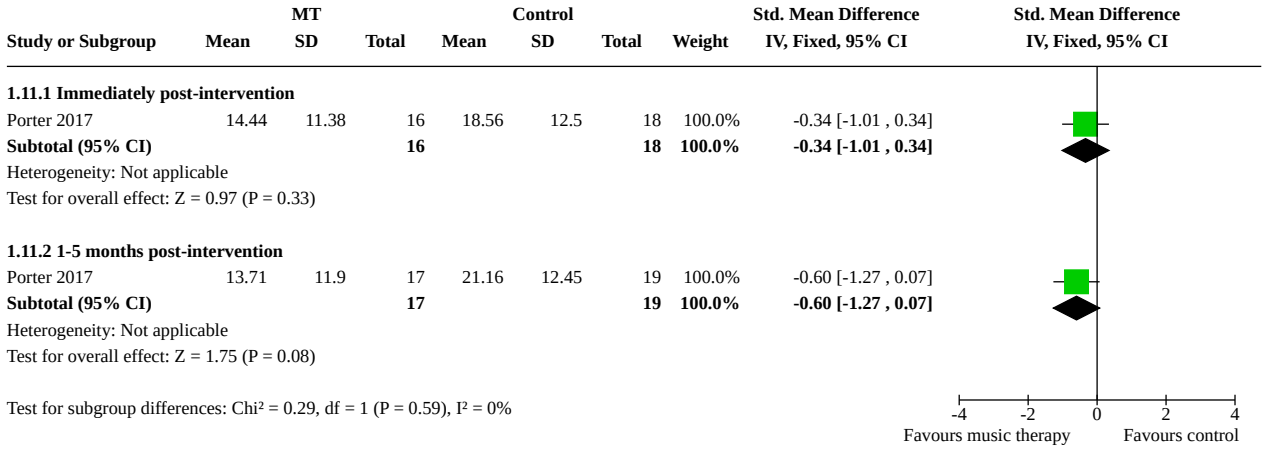
Study or Subgroup	MT Mean	MT SD	MT Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
1.9.1 Immediately post-intervention								
Kim 2008	16.4	1.43	10	15.8	1.55	10	36.5%	0.39 [-0.50 , 1.27]
Porter 2017	-2.1	0.26	6	-2.11	0.28	11	29.0%	0.03 [-0.96 , 1.03]
Thompson 2014	201.6	24.82	10	191.33	21.93	9	34.5%	0.42 [-0.50 , 1.33]
Subtotal (95% CI)			**26**			**30**	**100.0%**	**0.29 [-0.24 , 0.83]**
Heterogeneity: Chi² = 0.37, df = 2 (P = 0.83); I² = 0%								
Test for overall effect: Z = 1.08 (P = 0.28)								
1.9.2 1-5 months post-intervention								
Porter 2017	-2.03	0.23	6	-2.02	0.25	9	100.0%	-0.04 [-1.07 , 0.99]
Subtotal (95% CI)			**6**			**9**	**100.0%**	**-0.04 [-1.07 , 0.99]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.07 (P = 0.94)								
Test for subgroup differences: Chi² = 0.32, df = 1 (P = 0.57), I² = 0%								



Analysis 1.10. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 10: Identity formation



Analysis 1.11. Comparison 1: Music therapy vs placebo therapy or standard care, Outcome 11: Depression



ADDITIONAL TABLES

Table 1. Summarised characteristics of included studies

Category	Studies
Studies included in each version of this review	
First version (2006)	Brownell 2002; Buday 1995; Farmer 2003
Second version (2014)	Arezina 2011; Gattino 2011; Kim 2008; Lim 2010; Lim 2011; Thomas 2003; Thompson 2014 (which is a new report to a previously reported study)
Current update	Bharathi 2019; Bieleninik 2017 (with two more reports related to this study); Chen 2010; Chen 2013; Ghasemtabar 2015; Huang 2015; LaGasse 2014; Mateos-Moreno 2013; Moon 2010; Porter 2017 (with

Table 1. Summarised characteristics of included studies *(Continued)*

	another report related to this study); Rabeyron 2020; Sa 2020; Schwartzberg 2013; Schwartzberg 2016; Sharda 2018 (with another report related to this study); Yurteri 2019
Location	
North America	Canada: Sharda 2018 USA: Arezina 2011; Brownell 2002; Buday 1995; Farmer 2003; LaGasse 2014; Lim 2010; Lim 2011; Sa 2020 Schwartzberg 2013; Schwartzberg 2016; Thomas 2003
South America	Brazil: Gattino 2011
Asia	China: Chen 2010; Chen 2013; Huang 2015 Korea: Kim 2008; Moon 2010 India: Bharathi 2019 Iran: Ghasemtabar 2015
Europe	France: Rabeyron 2020 Spain: Mateos-Moreno 2013 Turkey: Yurteri 2019 UK: Porter 2017
Oceania	Australia: Thompson 2014
Multinational	Bieleninik 2017 (Australia, Austria, Brazil, Korea, Israel, Italy, Norway, UK, USA)
Design	
Parallel group	Bharathi 2019; Bieleninik 2017; Chen 2010; Chen 2013; Farmer 2003; Gattino 2011; Ghasemtabar 2015; Huang 2015; LaGasse 2014; Lim 2010; Mateos-Moreno 2013; Moon 2010; Porter 2017; Rabeyron 2020; Sa 2020; Schwartzberg 2013; Schwartzberg 2016; Sharda 2018; Thompson 2014; Yurteri 2019
Cross-over	Arezina 2011; Brownell 2002; Buday 1995; Kim 2008; Lim 2011; Thomas 2003
Individual participant data	
Available	Arezina 2011; Bharathi 2019; Bieleninik 2017; Brownell 2002; Farmer 2003; Gattino 2011; Kim 2008; LaGasse 2014; Porter 2017; Rabeyron 2020; Schwartzberg 2013; Schwartzberg 2016; Thomas 2003; Thompson 2014
Interventions	
Music therapy setting	Individual setting (one-to-one): Arezina 2011; Bieleninik 2017; Brownell 2002; Buday 1995; Farmer 2003; Gattino 2011; Kim 2008; Lim 2010; Lim 2011; Porter 2017; Sharda 2018; Thomas 2003; Yurteri 2019
	Group setting: Bharathi 2019; Ghasemtabar 2015; LaGasse 2014; Mateos-Moreno 2013; Rabeyron 2020; Sa 2020; Schwartzberg 2013; Schwartzberg 2016
	Either individually or in small groups: Yurteri 2019
	Family-based setting: Thompson 2014
	Unclear: Chen 2010; Chen 2013; Huang 2015; Moon 2010

Table 1. Summarised characteristics of included studies *(Continued)*

Music therapy frequency	Daily (for 1-2 weeks): Brownell 2002; Buday 1995; Farmer 2003; Lim 2010; Lim 2011; Schwartzberg 2013; Schwartzberg 2016
	Weekly: Arezina 2011; Gattino 2011; Kim 2008; Porter 2017; Rabeyron 2020; Sharda 2018; Thomas 2003; Thompson 2014
	Twice weekly: Chen 2013; Ghasemtabar 2015; LaGasse 2014; Mateos-Moreno 2013; Moon 2010; Yurteri 2019
	Several times a week: Bharathi 2019 (3 times a week); Chen 2010 (4 times a week); Huang 2015 (6 times a week)
	1 or 3 times a week: Bieleninik 2017
Music therapy content	Highly structured: Brownell 2002; Buday 1995; Chen 2010; Chen 2013; Farmer 2003; Lim 2010; Lim 2011; Moon 2010; Rabeyron 2020; Sa 2020; Schwartzberg 2013; Schwartzberg 2016
	Emphasis on interactive and relational aspects: Arezina 2011; Bharathi 2019; Bieleninik 2017; Gattino 2011; Ghasemtabar 2015; Huang 2015; Kim 2008; LaGasse 2014; Mateos-Moreno 2013; Porter 2017; Sharda 2018; Thomas 2003; Thompson 2014; Yurteri 2019
Comparators	
'Placebo' therapy	'Placebo' activity without music: Arezina 2011; Brownell 2002; Buday 1995; Farmer 2003; Kim 2008; LaGasse 2014; Lim 2010; Lim 2011; Moon 2010; Schwartzberg 2013; Schwartzberg 2016; Sharda 2018; Thomas 2003
	Passive music listening: Bharathi 2019; Rabeyron 2020
Standard care	Bieleninik 2017; Chen 2010; Chen 2013; Gattino 2011; Ghasemtabar 2015; Huang 2015; Mateos-Moreno 2013; Porter 2017; Sa 2020; Thompson 2014; Yurteri 2019
Outcomes	
Global improvement	Bieleninik 2017; Bharathi 2019; Kim 2008; LaGasse 2014; Porter 2017; Rabeyron 2020; Schwartzberg 2013; Thompson 2014
Social interaction	Arezina 2011; Bharathi 2019; Bieleninik 2017; Chen 2013; Gattino 2011; Ghasemtabar 2015; Kim 2008; LaGasse 2014; Porter 2017; Rabeyron 2020; Schwartzberg 2013; Sharda 2018; Thomas 2003; Thompson 2014
Non-verbal communication	Arezina 2011; Buday 1995; Chen 2013; Farmer 2003; Gattino 2011; Kim 2008; LaGasse 2014; Rabeyron 2020; Sharda 2018; Thomas 2003; Thompson 2014
Verbal communication	Buday 1995; Chen 2013; Farmer 2003; Gattino 2011; Lim 2010; Lim 2011; Rabeyron 2020; Schwartzberg 2013; Schwartzberg 2016; Sharda 2018; Thompson 2014
Quality of life	Bieleninik 2017; Sharda 2018; Yurteri 2019
Total autism symptom severity	Bharathi 2019; Bieleninik 2017; Chen 2010; Chen 2013; Huang 2015; LaGasse 2014; Mateos-Moreno 2013; Rabeyron 2020; Yurteri 2019
Adverse events	Bieleninik 2017; Porter 2017
Adaptive behaviour	Arezina 2011; Bieleninik 2017; Brownell 2002; Chen 2010; Kim 2008; Porter 2017; Rabeyron 2020; Sharda 2018; Thomas 2003

Table 1. Summarised characteristics of included studies *(Continued)*

Quality of family relationships	Kim 2008; Porter 2017; Thompson 2014
Identity formation	Moon 2010; Porter 2017
Depression	Porter 2017
Cognitive ability	Sa 2020

APPENDICES

Appendix 1. Search strategies 2020 onwards

When we ran the searches for this review update, we removed the search terms listed below from the original strategy because they did not identify any unique relevant records:

1. (speech adj3 disorder$).tw.
2. (language adj3 delay$).tw.

CENTRAL

Searched 7 July 2020. Limited to publication year 2013-2020 (217 records)
Searched 4 August 2021. Limited to records added between 7 July 2021 and 4 August 2021 (60 records)

#1 MeSH descriptor: [Music]
#2 MeSH descriptor: [Music Therapy]
#3 music*
#4 ((guided next imagery) near music)
#5 GIM
#6 vibroacoustic
#7 vibro-acoustic
#8 (sing or singing or song* or choral* or choir*)
#9 (percussion* or rhythm* or tempo* or melod*)
#10 improvis*
#11 (Nordoff-Robbin* or bonny*)
#12 ((auditory or acoustic or sound*) near/5 (stimulat* or cue*))
#13 (#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12)
#14 MeSH descriptor: [Child Development Disorders, Pervasive] 1 tree(s) exploded
#15 MeSH descriptor: [Neurodevelopmental Disorders] this term only
#16 MeSH descriptor: [Developmental Disabilities] this term only
#17 asperg* or autis* or kanner* or (childhood next schizophren*)
#18 ASD or ASDs or PDD or PDDs or PDD-NOS
#19 {or #14-#18}
#20 (#13 and #19) in Trials

MEDLINE Ovid

Searched 8 July 2020 (217 records)
Searched 4 August 2021 (29 records)

1 music therapy/
2 music/
3 music$.tw,kf.
4 ((guided imagery adj3 music) or gim).tw,kf.
5 (vibro-acoustic$ or vibroacoustic$).tw,kf.
6 (sing or singing or song$ or choral$ or choir$).tw,kf.
7 (percussion$ or rhythm$ or tempo).tw,kf.
8 melod$.tw,kf.
9 improvis$.tw,kf.

10 (Nordoff-Robbin$ or bonny$).tw,kf.
11 ((auditory or acoustic or sound$) adj5 (stimulat$ or cue$)).tw,kf.
12 or/1-11
13 exp child development disorders, pervasive/
14 Developmental Disabilities/
15 Neurodevelopmental Disorders/
16 pervasive development$ disorder$.tw,kf.
17 (pervasive adj3 child$).tw,kf.
18 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw,kf.
19 autis$.tw,kf.
20 asperg$.tw,kf.
21 kanner$.tw,kf.
22 childhood schizophreni$.tw,kf.
23 or/13-22
24 randomized controlled trial.pt.
25 controlled clinical trial.pt.
26 randomi#ed.ab.
27 placebo$.ab.
28 drug therapy.fs.
29 randomly.ab.
30 trial.ab.
31 groups.ab.
32 or/24-31
33 exp animals/ not humans.sh.
34 32 not 33
35 12 and 23 and 34
36 limit 35 to ed=20130701-20200626 Annotation: Final line 2020 update search
37 (20200626* or 20200627* or 20200628* or 20200629* or 20200630* or 202007* or 202008* or 202009* or 202010* or 202011* or 202012* or 2021*).dt,ez,da.
38 35 and 37 Annotation: Final line 2021 top-up search

MEDLINE In-Process & Other Non-Indexed Citations Ovid

Searched 6 July 2020 (84 records)
Searched 4 August 2021 (10 records)

1 music$.tw,kf.
2 (guided imagery adj3 music).tw,kf. or gim.tw,kf.
3 (vibro-acoustic$ or vibroacoustic$).tw,kf.
4 (sing or singing or song$ or choral$ or choir$).tw,kf.
5 (percussion$ or rhythm$ or tempo).tw,kf.
6 melod$.tw,kf.
7 improvis$.tw,kf.
8 (Nordoff-Robbin$ or bonny$).tw,kf.
9 ((auditory or acoustic or sound$) adj5 (stimulat$ or cue$)).tw,kf.
10 or/1-9
11 pervasive development$ disorder$.tw,kf.
12 (pervasive adj3 child$).tw,kf.
13 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw,kf.
14 autis$.tw,kf.
15 asperg$.tw,kf.
16 kanner$.tw,kf.
17 childhood schizophreni$.tw,kf.
18 or/11-17
19 10 and 18
20 (random$ or trial$ or control$ or group$ or placebo$ or blind$ or prospectiv$ or longitudinal$ or meta-analys$ or systematic review $).tw,kf.
21 19 and 20

MEDLINE Epub Ahead of Print Ovid

Searched 6 July 2020 (14 records)
Searched 4 August 2021 (14 records)

1 music$.tw,kf.
2 (guided imagery adj3 music).tw,kf. or gim.tw,kf.
3 (vibro-acoustic$ or vibroacoustic$).tw,kf.
4 (sing or singing or song$ or choral$ or choir$).tw,kf.
5 (percussion$ or rhythm$ or tempo).tw,kf.
6 melod$.tw,kf.
7 improvis$.tw,kf.
8 (Nordoff-Robbin$ or bonny$).tw,kf.
9 ((auditory or acoustic or sound$) adj5 (stimulat$ or cue$)).tw,kf.
10 or/1-9
11 pervasive development$ disorder$.tw,kf.
12 (pervasive adj3 child$).tw,kf.
13 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw,kf.
14 autis$.tw,kf.
15 asperg$.tw,kf.
16 kanner$.tw,kf.
17 childhood schizophreni$.tw,kf.
18 or/11-17
19 10 and 18
20 (random$ or trial$ or control$ or group$ or placebo$ or blind$ or prospectiv$ or longitudinal$ or meta-analys$ or systematic review $).tw,kf.
21 19 and 20

Embase Ovid

Searched 8 July 2020 (168 records)
Searched 4 August 2021 (38 records)

1 exp music/
2 music therapy/
3 music$.tw,kw.
4 (guided imagery adj3 music).tw,kw.
5 GIM.tw,kw.
6 (vibro-acoustic therapy or vibroacoustic therapy).tw,kw.
7 (sing or singing or song$ or choral$ or choir$).tw,kw.
8 (percussion$ or rhythm$).tw,kw.
9 melod$.tw,kw.
10 improvis$.tw,kw.
11 (Nordoff-Robbin$ or bonny$).tw,kw.
12 ((auditory or acoustic or sound$) adj5 (stimulat$ or cue$)).tw,kw.
13 or/1-12
14 exp autism/
15 developmental disorder/
16 pervasive development$ disorder$.tw.
17 (PDD or PDDs or PDD_NOS or ASD or ASDs).tw,kw.
18 autis$.tw.
19 asperg$.tw.
20 kanner$.tw.
21 childhood schizophreni$.tw.
22 or/14-21
23 13 and 22
24 Randomized controlled trial/
25 Controlled clinical study/
26 random$.ti,ab.
27 randomization/
28 intermethod comparison/
29 placebo.ti,ab.
30 (compare or compared or comparison).ti.
31 ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab.
32 (open adj label).ti,ab.
33 ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab.
34 double blind procedure/
35 parallel group$1.ti,ab.

36 (crossover or cross over).ti,ab.
37 ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant $1)).ti,ab.
38 (assigned or allocated).ti,ab.
39 (controlled adj7 (study or design or trial)).ti,ab.
40 (volunteer or volunteers).ti,ab.
41 human experiment/
42 trial.ti.
43 or/24-42
44 (random$ adj sampl$ adj7 ("cross section$" or questionnaire$1 or survey$ or database$1)).ti,ab. not (comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.)
45 Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.) (238532)
46 (((case adj control$) and random$) not randomi?ed controlled).ti,ab.
47 (Systematic review not (trial or study)).ti.
48 (nonrandom$ not random$).ti,ab.
49 "Random field$".ti,ab.
50 (random cluster adj3 sampl$).ti,ab.
51 (review.ab. and review.pt.) not trial.ti.
52 "we searched".ab. and (review.ti. or review.pt.)
53 "update review".ab.
54 (databases adj4 searched).ab.
55 (rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment/
56 Animal experiment/ not (human experiment/ or human/)
57 or/44-56
58 43 not 57
59 23 and 58
60 limit 59 to yr="2013 -Current"
61 limit 59 to yr="2020 -Current"

LILACS

Searched 7 July 2020. Limited by year =2013-2020 (7 records)
Searched 4 August 2021. Limited by year =2020-2021 (0 records)

(tw:((music* OR gim OR percussion* OR rhythm* OR tempo OR improvis* OR melod* OR sing OR singing OR song* OR choral* OR choir* OR auditory OR acoustic OR sound* OR vibro*))) AND (tw:((autis* OR asperger* OR "pervasive developmental" OR ASD OR PDD OR PDD-nos)))

Results filtered by source (LILACS) and study type (controlled clinical trial)

APA PsycInfo Ovid

Searched 6 July 2020 (91 records)
Searched 4 August 2021 (15 records)

1 exp music/
2 music therapy/
3 music$.tw.
4 (guided imag$ adj3 music* or gim).tw.
5 GIM.tw.
6 (vibroacoustic$ or vibro-acoustic$).tw.
7 rhythm/ or tempo/
8 (percussion$ or rhythm$ or tempo).tw.
9 singing/
10 (sing or singing or song$ or choral$ or choir$).tw.
11 melod$.tw.
12 improvis$.tw.
13 (Bonny or Nordoff$).tw.
14 ((auditory or acoustic or sound$) adj5 (stimulat$ or cue$)).tw.
15 or/1-14
16 exp autism spectrum disorders/
17 Developmental Disabilities/
18 neurodevelopmental disorders/

19 pervasive development$ disorder$.tw.
20 (PDD or PDDs or PDD-NOS or ASD or ASDs).tw.
21 autis$.tw.
22 asperger$.tw.
23 kanner$.tw.
24 childhood schizophreni$.tw.
25 or/16-24
26 randomized controlled trials/
27 randomized clinical trials/
28 Clinical Trials/
29 exp treatment effectiveness evaluation/
30 Placebo/
31 (control$ adj3 (study or studies or trial$ or group$)).tw.
32 (allocat$ or assign$).ab.
33 (placebo or treatment as usual or tau).ab.
34 (random$ or RCT).tw.
35 ((singl$ or doubl$ or tripl$ or trebl$) adj3 (blind$ or mask$)).tw.
36 (crossover$ or cross-over$).tw.
37 ((evaluat$ or effectiveness$) adj3 (study or studies or research$)).tw.
38 or/26-37
39 15 and 25 and 38
40 limit 39 to up=20130701-20200629
41 limit 39 to yr="2013 -Current"
42 40 or 41
43 limit 39 to up=20200629-20210726
44 limit 39 to yr="2020 -Current"
45 43 or 44

CINAHL Plus EBSCOhost

Searched 6 July 2020 (75 records)
Searched 4 August 2021 (15 records)

S1 (MH "Music")
S2 (MH "Music Therapy")
S3 music*
S4 ((guided imagery N3 music) or gim)
S5 (vibro-acoustic* or vibroacoustic*)
S6 (Nordoff* or Bonny*)
S7 (percussion* or rhythm* or tempo)
S8 melod*
S9 (MH "Singing")
S10 (sing or singing or song* or choral* or choir*)
S11 (Nordoff* or Bonny*)
S12 ((auditory or acoustic or sound*) N5 (stimulat* or cue*))
S13 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12
S14 (MH "Child Development Disorders, Pervasive+")
S15 (pervasive development* disorder* or PDD or PDDs)
S16 (autis* or ASD or ASDs)
S17 Asperger*
S18 Kanner*
S19 childhood schizophren*
S20 (MH "Developmental Disabilities")
S21 S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20
S22 S13 AND S21
S23 MH randomized controlled trials
S24 MH double-blind studies
S25 MH single-blind studies
S26 MH random assignment
S27 MH pretest-posttest design
S28 MH cluster sample5
S29 TI (randomised OR randomized)
S30 AB (random*)

S31 TI (trial)
S32 MH (sample size) AND AB (assigned OR allocated OR control)
S33 MH (placebos)
S34 PT (randomized controlled trial)
S35 AB (control W5 group)
S36 MH (crossover design) OR MH (comparative studies)
S37 AB (cluster W3 RCT)
S38 MH animals+
S39 MH (animal studies)
S40 TI (animal model*)
S41 S38 OR S39 OR S40
S42 MH (human)
S43 S41 not S42
S44 S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37
S45 S44 not S43
S46 S22 AND S45
S47 EM 20130701-
S48 S46 AND S47
S49 EM 20200701-
S50 S46 AND S49

ERIC EBSCOhost

Searched 6 July 2020. Limited by year =2013-2020 (58 records)
Searched 4 August 2021. Limited by year =2020-2021 (19 records)

S1 DE "Developmental Disabilities"
S2 DE "Pervasive Developmental Disorders" OR DE "Asperger Syndrome" OR DE "Autism"
S3 (pervasive development* disorder* or PDD or PDDs)
S4 (autis* or ASD or ASDs)
S5 Asperger*
S6 Kanner*
S7 childhood schizophren*
S8 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7
S9 DE "Music Therapy"
S10 (DE "Music" OR DE "Music Activities")
S11 music*
S12 ((guided imagery N3 music) or gim)
S13 (vibro-acoustic* or vibroacoustic*)
S14 (Nordoff* or Bonny*)
S15 (percussion* or rhythm* or tempo)
S16 melod*
S17 (DE "Singing")
S18 (sing or singing or song* or choral* or choir*)
S19 improvis*
S20 ((auditory or acoustic or sound*) N5 (stimulat* or cue*))
S21 S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20
S22 S8 AND S21
S23 DE "Randomized Controlled Trials" OR DE "Meta Analysis" OR DE "Evaluation Research" OR DE "Control Groups" OR DE "Experimental Groups" OR DE "Longitudinal Studies" OR DE "Followup Studies" OR DE "Program Effectiveness" OR DE "Program Evaluation"
S24 TI (random* or trial* or experiment* or prospectiv* OR longitudinal or BLIND* or CONTROL*) OR AB (random* or trial* or experiment* or prospectiv* OR longitudinal or BLIND* or CONTROL*)
S25 S23 OR S24
S26 S22 AND S25

Sociological Abstracts Proquest

Searched 4 August 2021. Limited by year =2013-2020 (9 records)

((su.EXACT("Music") or NOFT(music* or guided imag* or GIM or vibro-acoustic therapy* or vibroacoustic therapy* or Bonny* or Nordoff* or singing or song* or choral* or choir* or percussion* or rhythm* or improvis*) OR NOFT((auditory or acoustic or sound*) near/5 (stimulat* or cue*))) and (su.EXACT("autism") or NOFT(autism* or asperg* or "pervasive development* disorder*" or ("childhood schizophrenia") or Kanner*)) and NOFT(random* or placebo* or trial* or blind* or group* or control or controlled or RCT or TAU or "usual treatment" or "treatment as usual")

Proquest Global Dissertations & Theses

Searched 4 August 2021. Limited by year =2013-2021 (9 records)

(NOFT(music* or guided imag* or GIM or vibro-acoustic therapy* or "vibroacoustic therapy*" or Bonny* or Nordoff* or singing or song* or choral* or choir* or percussion* or rhythm* or improvis*) OR NOFT((auditory or acoustic or sound*) near/5 (stimulat* or cue*))) and (NOFT(autism* or asperg* or "pervasive development* disorder*" or "childhood schizophrenia*" or Kanner*)) and NOFT((random* or placebo* or trial* or blind* or group* or control or controlled or RCT or TAU or "usual treatment" or "treatment as usual"))

Proquest Music Periodicals Database

Searched 6 July 2020 No limits (138 records)
Searched 4 August 2021. Limited by year =2020-2021 (7 records)

(NOFT(MUSIC* n/3 THERAP*) OR NOFT((guided imagery N/3 music) OR gim) OR NOFT(percussion* OR rhythm* OR tempo OR improvis* OR melod* OR sing OR singing OR song* OR choral* OR choir*) OR NOFT((auditory OR acoustic OR sound*) N/5 (stimulat* OR cue*)) OR NOFT(Nordoff* OR Bonny*) OR NOFT((vibro-acoustic* OR vibroacoustic*))) AND NOFT(autis* OR asperger* OR kanner* OR "pervasive developmental" OR ASD OR ASDs OR PDD or PDDs OR PDD-NOS) AND NOFT(random* OR trial* OR experiment* OR prospectiv* OR longitudinal OR blind* OR control* OR placebo OR "treatment as usual" OR TAU OR intervention* OR treat*)

Proquest Performing Arts Periodicals Database

Searched 6 July 2020 No limits (0 records)
Searched 4 August 2021. Limited by year =2020-2021 (1 record)

(NOFT((MUSIC* n/3 THERAP*) OR (guided imagery N/3 music) OR gim OR percussion* or rhythm* or tempo or improvis* OR melod* OR sing OR singing OR song* OR choral* OR choir*) OR NOFT((auditory OR acoustic OR sound*) N/5 (stimulat* or cue*)) OR NOFT(Nordoff* OR Bonny*) OR NOFT((vibro-acoustic* OR vibroacoustic*)) OR NOFT((MUSIC* n/3 THERAP*) OR ("guided imagery" N/3 music) OR gim OR percussion* OR rhythm* OR tempo OR improvis* OR melod* OR sing OR singing OR song* OR choral* OR choir*)) AND NOFT(autis* OR asperger* OR kanner* OR "pervasive developmental" OR ASD OR ASDs OR PDD OR PDDs OR PDD-NOS)

RILM Abstracts of Music Literature Online EBSCOhost

Searched 6 July 2020 No limits (95 records)
Searched 4 August 2021. Limited by year =2020-2021 (1 record)

S1 SU "therapy—music therapy --*"
S2 music* N3 therap*
Database - RILM Abstracts of Music Literature (1967 to present)
S3 ((guided imagery N3 music) or gim)
Database - RILM Abstracts of Music Literature (1967 to present)
S4 (vibro-acoustic* or vibroacoustic*)
Database - RILM Abstracts of Music Literature (1967 to present)
S5 (Nordoff* or Bonny*)
Database - RILM Abstracts of Music Literature (1967 to present)
S6 (percussion* or rhythm* or tempo or improvis*)
Database - RILM Abstracts of Music Literature (1967 to present)
S7 melod*
Database - RILM Abstracts of Music Literature (1967 to present)
S8 (sing or singing or song* or choral* or choir*)
Database - RILM Abstracts of Music Literature (1967 to present)
S9 ((auditory or acoustic or sound*) N5 (stimulat* or cue*))
Database - RILM Abstracts of Music Literature (1967 to present)
S10 S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9
Database - RILM Abstracts of Music Literature (1967 to present)
S11 (autis* or ASD or ASDs)
Database - RILM Abstracts of Music Literature (1967 to present)
S12 Asperger*
Database - RILM Abstracts of Music Literature (1967 to present)
S13 Kanner*
Database - RILM Abstracts of Music Literature (1967 to present)
S14 childhood schizophren*
Database - RILM Abstracts of Music Literature (1967 to present)
S15 (pervasive development* disorder* or PDD or PDDs or PDD-NOS)
Database - RILM Abstracts of Music Literature (1967 to present)

S16 S11 OR S12 OR S13 OR S14 OR S15
Database - RILM Abstracts of Music Literature (1967 to present)
S17 S10 AND S16
Database - RILM Abstracts of Music Literature (1967 to present)
S18 (random* or trial* or experiment* or prospectiv* OR longitudinal or blind* or control* or placebo or "treatment as usual" or TAU)
Database - RILM Abstracts of Music Literature (1967 to present)
S19 S17 AND S18
Database - RILM Abstracts of Music Literature (1967 to present)

Cochrane Database of Systematic Reviews

Searched 7 July 2020 (9 records)
Searched 4 August 2021 (0 records)

#1 MeSH descriptor: [Music] this term only
#2 MeSH descriptor: [Music Therapy] this term only
#3 music*:ti,ab,kw
#4 (guided next imagery):ti,ab,kw
#5 GIM:ti,ab,kw
#6 vibroacoustic:ti,ab,kw
#7 vibro-acoustic:ti,ab,kw
#8 (sing or singing or song* or choral* or choir*):ti,ab,kw
#9 ((percussion* or rhythm* or tempo* or melod*) and music*):ti,ab,kw
#10 (improvis* near music*):ti,ab,kw
#11 (Nordoff-Robbin* or bonny*):ti,ab,kw
#12 ((auditory or acoustic or sound*) near/5 (stimulat* or cue*)):ti,ab,kw
#13 (#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12)
#14 MeSH descriptor: [Child Development Disorders, Pervasive] this term only
#15 MeSH descriptor: [Neurodevelopmental Disorders] this term only
#16 MeSH descriptor: [Developmental Disabilities] this term only
#17 asperg* or autis* or kanner* or (childhood next schizophren*):ti,ab,kw
#18 (ASD or ASDs or PDD or PDDs or PDD-NOS):ti,ab,kw
#19 {or #14-#18}
#20 (#13 and #19) with Cochrane Library publication date Between Jul 2013 and Jul 2020, in Cochrane Reviews, Cochrane Protocols
#21 (#13 and #19) with Cochrane Library publication date Between Jul 2020 and Aug 2021, in Cochrane Reviews.Cochrane Protocols

Epistemonikos

Searched 7 July 2020 (22 records)
Searched 4 August 2021. Limited to records added from 7 July 2020 to 4 August 2021 (2 records)

title:((autis* OR asperger* OR "pervasive developmental" OR ASD OR PDD OR PDD-nos)) AND (title:(((cue* OR signal*) AND (auditor* OR acoustic* OR sound))) OR abstract:(((cue* OR signal*) AND (auditor* OR acoustic* OR sound))))

title:((autis* OR asperger* OR "pervasive developmental" OR ASD OR PDD OR PDD-nos)) AND (title:(music*) OR abstract:(music*))

title:((autis* OR asperger* OR "pervasive developmental" OR ASD OR PDD OR PDD-nos)) AND (title:(sing OR singing OR choral OR choir OR tempo OR improvis* OR rhythm) OR abstract:(sing OR singing OR choral OR choir OR tempo OR improvis* OR rhythm))

ClinicalTrials.gov

Searched 7 July 2020 (22 records)
Searched 5 August 2021. Limited to records first posted between 7 July 2020 and 5 August 2021 (7 records)

Advanced search: Interventional Studies | Autism OR autistic OR asperger OR ASD OR pervasive developmental disorder OR PDD OR PDD-NOS | music OR tempo OR rhythm OR GIM or guided imagery OR acoustic OR auditory OR sound

WHO ICTRP

Search attempted 7 July 2020, but due to heavy traffic generated by the COVID-19 outbreak, the ICTRP Search Portal was not responding from outside WHO.

Searched 5 August 2021 No limits (21 records)

Standard search: (music AND autism) OR (music AND ASD) OR (music AND asperger)

Appendix 2. Criteria for assigning risk of bias judgements

Random sequence generation

We judged the risk of bias for random sequence generation as follows.

1. Studies were judged to be at low risk of bias if participants were allocated to treatment interventions using randomisation, such as computer-generated random numbers, a random numbers table, or coin-tossing.
2. Studies were judged to be at unclear risk of bias if the randomisation method was not clearly stated or was unknown.
3. Studies were judged to be at high risk of bias if the method sequence generation was non-random.

Randomised as well as controlled clinical trials were included in the review, as noted above (see Types of studies).

Allocation concealment

We judged the risk of bias for allocation concealment as follows.

1. Studies were judged to be at low risk of bias if the allocation concealment was adequate; participants and researchers were unaware of participants' future allocation to an intervention until after decisions about eligibility were made, and informed consent was obtained.
2. Studies were judged to be at unclear risk of bias if the methods used for allocation concealment were not described in detail.
3. Studies were judged to be at high risk of bias if the allocation concealment was inadequate; allocation was not concealed, either from participants before informed consent or from researchers before decisions about inclusion were made (this will always be the case for quasi-randomised studies).

Both randomised and controlled trials were judged using the same criteria for gaining a descriptive measure of study quality.

Blindness of participants and personnel

Due to the nature of the intervention, it was not possible to blind those who delivered music therapy or those who received it. Consequently, neither participants nor therapists of the studies under review could be declared to be blinded. However, although autistic individuals were not blinded, this was unlikely to introduce bias as usually they are not fully aware of available treatment options or study design (Cheuk 2011). The possible risk of bias introduced by therapists administering the intervention was unknown. Therefore, we judged the risk of performance bias as unclear in all studies included in this review.

Blinding of outcome assessors

We determined whether those who assessed and coded the outcome measures were blind to treatment assignment using the following categories.

1. Studies were judged to be at low risk of bias if the assessor was blind to treatment assignment.
2. Studies were judged to be at unclear risk of bias if the blinding of the assessor was not reported and information was not available from the researchers.
3. Studies were judged to be at high risk of bias if the assessor was not blind to treatment assignment.

Completeness of outcome data

We assessed whether study authors adequately dealt with missing data as follows.

- Studies were judged to be at low risk of bias if the number of participants randomised to groups was clear and it was clear that all participants completed the trials in all participant groups. Studies were also judged to be at low risk of bias if outcome data were missing in both intervention groups, but reasons for these were both reported and balanced across groups.
- Studies were judged to be at unclear risk of bias if information about which participants completed the study could not be acquired by contacting the study authors.
- Studies were judged to be at high risk of bias if there was clear evidence of attrition or exclusion from analysis in at least one participant group that was likely to be related to the true outcome.

Selective reporting

We judged the risk of selective outcome reporting as follows.

1. Studies were judged to be at low risk of bias if all collected data seemed to be reported and all expected outcomes were reported.
2. Studies were judged to be at unclear risk of bias if it was not clear whether other data were collected and not reported.
3. Studies were judged to be at high risk of bias if data for one or more expected outcomes were missing.

Other bias

Through assessment, we determined whether any other bias was present in the trials, including inadequate music therapy methods (not corresponding to the definition of music therapy used for this review) or inadequate music therapy training of therapists delivering the intervention (without formal training based on the particular country's regulations for music therapy).

WHAT'S NEW

Date	Event	Description
16 November 2021	New search has been performed	A top-up search for new studies was conducted, resulting in the inclusion of one new study in the qualitative synthesis and three new ongoing studies.
9 April 2021	New search has been performed	A search for new studies was conducted, resulting in the inclusion of fifteen new studies; the categories of outcome measures were adapted; based on the added studies' findings, new meta-analyses were performed, and pre-existing results and conclusions were modified.
16 October 2020	New citation required and conclusions have changed	Updated review with one new author.

HISTORY

Protocol first published: Issue 3, 2003
Review first published: Issue 2, 2006

Date	Event	Description
18 March 2016	Amended	Abstract, main results - adding the word 'therapy' in the second sentence
2 December 2013	New search has been performed	A search for new studies was conducted, resulting in the inclusion of seven new studies; based on the added studies' findings, the categories of outcome measures were revised, new meta-analyses were performed, and pre-existing results and conclusions were modified.
31 March 2013	New citation required and conclusions have changed	Updated review with two new authors.
5 November 2009	Amended	Minor edit in background.
10 November 2008	Amended	Converted to new review format.
21 February 2006	Amended	Minor update
29 January 2006	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

CG is the guarantor, conceived the review, designed the protocol and co-ordinated the reviewing. MG co-ordinated this review's update. CE, LFP, and MG searched for studies. CE, LFP, MG, and GV screened search results. LFP, MG, KM, and GV assessed risk of bias and assessed

the certainty in the body of evidence. CE, LFP, MG, CG, KM, and GV extracted data, analysed data, interpreted data, wrote the report, and approved the full review.

Contribution of previous authors: Tony Wigram, co-author of the 2006 version of this review, contributed to the development of the protocol, extracted and analysed data, and helped with writing the original report.

DECLARATIONS OF INTEREST

Cochavit Elefant (CE), Monika Geretsegger (MG), Christian Gold (CG), and Karin Mössler (KM) are clinically trained music therapists. CE, MG and KM report having been involved in publications from one study included in this review (Bieleninik 2017), without it supporting or influencing their work on this review; the study was funded by the Research Council of Norway (grant 213844, the Clinical Research and Mental Health Programmes); POLYFON Knowledge Cluster for Music Therapy; The Grieg Academy Department of Music, University of Bergen, Norway; and a range of further governmental and university funding sources and foundations across participating countries (see Characteristics of included studies for details). Assessment of eligibility, extraction of data, and assessment of risk of bias and the certainty of the evidence of this study was performed by two independent review authors who were not involved in the study.

Cochavit Elefant (CE) has declared that she has no other conflicts of interest.

Monika Geretsegger (MG) has declared that she has no other conflicts of interest.

Christian Gold (CG) is an Associate Editor of the Cochrane Developmental, Psychosocial and Learning Problems Review Group, without it supporting or influencing his work on this review. He is entirely excluded from the editorial decisions and related activities concerning this review. CG is a member of the Austrian Professional Association of Music Therapists, and until 2020, worked as a health professional. CG reports a grant from the Kavli Trust for the project 'Music for Autism' (M4A), paid to NORCE Norwegian Research Centre. CG reports being involved in publications from three studies included in this review (Bieleninik 2017; Kim 2008; and Thompson 2014). He also reports publishing an invited commentary in the Lancet Child and Adolescent Health in 2019. CG reports being the PI of two studies (Bieleninik 2017, funded by the Research Council of Norway, National Institute of Health Research; and NCT04936048, funded by the Kavli Trust), as well as a statistical advisor for one study (Kim 2008, funded by Aalborg University) eligible for inclusion in this review; none of which supported or influenced his work on this review. Assessment of eligibility, extraction of data, and assessment of risk of bias and the certainty of the evidence of these three studies were performed by two independent review authors who were not involved in the studies.

Laura Fusar-Poli is a Clinical Research Fellow (RTD-A) and psychiatrist in training at the University of Catania, Italy. She has declared that she has no conflicts of interest.

Giovanni Vitale is a child psychiatrist in training. He has declared that he has no conflicts of interest.

Karin Mössler has declared that she has no other conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- NORCE Norwegian Research Centre, Bergen, Norway

  Monika Geretsegger, Christian Gold, and Karin A Mössler received salary support from NORCE during this update.
- University of Vienna, Austria

  Christian Gold received salary support from the University of Vienna during this update.
- University of Haifa, Israel

  Cochavit Elefant received salary support from the University of Haifa during this update.

External sources

- Kavli Trust, Norway

  Kavli Trust funded the Project 'Music for Autism (M4A)'. This review update is included among the expected publications from M4A. Kavli Trust had no role in the design, conduct or publication of this review update.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Differences between protocol and original review

In compliance with the developments in systematic review methods since publication of the first version of this review (Gold 2006), a distinction was made between primary and secondary outcome measures, and risk of bias and summary of findings tables were included in the 2014 update (Geretsegger 2014).

Differences between previous versions of the review and this update

Title

We have updated the term 'people with autism spectrum disorder' to 'autistic people' throughout the text to meet the preferences generally expressed by autistic people.

Types of outcome measures

The knowledge of the condition itself and its nosological classification have changed a lot since the protocol was published in 2003. Back then, there was still the conception that ASD was mostly a paediatric condition; the concept of ASD was introduced in 2013 with DSM-5 (APA 2013), and now we know that ASD is a life-long condition which results in different aspects being relevant (i.e. mental health problems or self-esteem). To ensure that all user-important outcomes are addressed (McKenzie 2021), and to update our approach in correspondence with changes that occurred in the knowledge about ASD in recent years, we adapted the outcome categories used in the previous versions of the review as described in the Methods section (see Types of outcome measures). We also merged some previously separate outcome measures to broader outcome categories to keep the review focused and manageable for users.

Electronic searches

We made the following changes to the databases listed in the protocol (Gold 2003) and used in the in the previous versions of the review (Geretsegger 2014; Gold 2006), due to changes to standard search methods at the Cochrane Developmental, Psychosocial and Learning Problems Review Group. For this update, we searched the *Cochrane Database of Systematic Reviews* (CDSR), two trials registers, and two daily updated segments of MEDLINE (MEDLINE In-Process and Other Non-Indexed Citations, and MEDLINE EPub Ahead of Print). We also added Epistemonikos as a source of reviews.

Two subscription databases used in previous versions of the review (Dissertation Abstracts and ASSIA) were no longer available for the updated review, and were replaced by Proquest Global Dissertations & Theses, and three specialist music databases (Proquest Music Periodical Database, Proquest Performing Arts Periodicals Database and RILM Abstracts of Music Literature Online).

Three other resources used in previous versions of the review were not used for this update as they were no longer maintained (musictherapyworld.net website) or no longer updated (Music Therapy Research CD ROM, AMTA 1999; Music Therapy World Info-CD ROM IV, Aldridge 2002).

Data collection and analysis

In this update, we applied the Cochrane's Screen4Me workflow to help assess the search results. We also added a new section, 'Summary of findings and assessment of the certainty of the evidence', in line with changes to standard methods in Cochrane.

Appendix 1

We removed search terms that identified irrelevant studies in the original search.

INDEX TERMS

Medical Subject Headings (MeSH)

*Autistic Disorder [therapy]; Bias; *Music Therapy; Odds Ratio; Quality of Life

MeSH check words

Adolescent; Child; Humans